A STANDARDISATION OF THE BOEHM TEST OF BASIC CONCEPTS ON CHILDREN IN ENGLAND.

A thesis submitted to the University of Manchester for the degree of Ph.D. in the Faculty of Education.

1984

Eileen F. Smith

THE UNIVERSITY OF MANCHESTER

ABSTRACT OF THESIS submitted by Eileen F. Smith

for the Degree of Doctor of Philosophy and entitled

A standardisation of the Boehm Test of Basic Concepts on children in England.

Date of submission May 1984

ABSTRACT

The Boehm Test of Basic Concepts (BTBC) Form A, previously only available in an American version with US norms, was standardised on children in England (N = 1928). The test directions were modified for British use. A table is presented for converting raw scores to standard scores (T scores) by age level in half-years from 3:6 to 7:6. An equivalence study found the two alternate forms of the BTBC (Forms A and B) to yield roughly equivalent total scores. Two short forms of the BTBC Form A were compiled and separately scaled for children aged 3½ to 5 years.

Content validity evidence suggests that most children will encounter the 50 BTBC concepts before their eighth birthday. Most, but not all, of these concepts were found to be seminal to the infant school curriculum.

Predictive validity evidence shows relationships between BTBC scores and achievements in language, mathematics and reading as assessed by teachers.

The BTBC is considered to be a multidimensional test of general cognitive development as well as being a guide to concept acquisition.

Non-linguistic information and limitations in item design, as well as semantic, syntactical and conceptual knowledge, influence children's responses to the BTBC items. Diversity in the efficacy of individual test items is apparent.

Relationships were found between the BTBC scores and both social class and types of domiciliary areas.

See Notes on reverse

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

This research is dedicated with affection and respect
to the memory of my first mentors in Educational Psychology,
Professor Stephen Wiseman, Professor Frank Warburton and Dr. Winifred Langan,
and to the first supervisor of this research,
the late Dr. Tom Fitzpatrick.

Acknowledgements

I should like to express gratitude to

Dr. C.D. Elliott for his patience and scrupulous supervision;

Dr. Anne E. Boehm, for her encouraging interest and for sending unpublished material on her test;

The Psychological Corporation, New York for permission to Anglicise certain words in the test instructions;

The Chief Education Officers, Administrative Officers and Advisors of Barnsley, Doncaster and Humberside, for permission to use schools and for supplying information;

The headteachers, class teachers and ancillary staff of all the schools visited, for their cooperation and generous hospitality;

A.N. James, Head of Statistical Services, NFER, for advice on sampling;

H.M. Inspectorate for help with the classification of schools;

Mr. D.C. Gill, Director of Planning, Humberside County Council, Mr. Haslam, Research Officer, Doncaster M.B.C., Mr. Hart and Mr. Welsh, Town Planning Officers, Doncaster M.B.C., for information on populations;

Eric Wilkinson, Director of the Strathclyde Experiment, for sending copies of unpublished reports;

Peter Hopwood and the staff of High Melton library, D.M.I.H.E.;

Fred Harmer and Jean Melrose, for advice on curriculum materials;

Doug Anderson and Ella Williams, for advice on demography;

Caroline for her excellent typing;

Seán (my instant thesaurus), Mollie and the many friends who helped with their support, tolerance, hospitality and the occasional, instant, old-fashioned grammar lesson;

Edna, for her lesson in logic, and Bill, for his physics lesson.

CONTENTS

Appendices

List of Tables

* There is no Table 31

List of Figures

-oOo-

In order to avoid clumsiness in the text, I have used the pronoun "he" when referring to either gender. This is due to inadequacies in the English language and not to my own sexual discrimination.

CHAPTER 1

An Outline of the Research

The Boehm Test of Basic Concepts (BTBC) was designed to measure children's mastery of concepts considered necessary for achievement in the first years of school (Boehm, 1971). Boehm queried the assumption that children have mastered the basic concepts necessary for understanding and following directions by the time of school entry (ibid.).

Use of the BTBC in the UK has been limited by the lack of British norms. The BTBC manual provides only percentile equivalents of raw scores by US school grades and socioeconomic level.

The BTBC (slightly modified) was standardised on children in England. A representative sample of 1928 children, aged 3½ to 7½ years, in 78 schools, was tested with Form A by individual administration. The mean reliability estimate was .81 (KR_{20}) and the mean raw score standard error of measurement was .27. No overall sex differences were found. Raw scores were converted to T scores, normalised on percentiles, and these standard scores (after smoothing) are presented in a conversion table by age level in half-years from 3:6 to 7:6. A ceiling effect is apparent at the upper end of the age range.

An equivalence study of the two alternate forms of the BTBC, Form A and Form B, yielded a mean correlation coefficient of .80 between the two forms. Subjects were 144 children aged 4 to 7 years. The results confirm Boehm's statement that "the two forms yield essentially equivalent scores" (Boehm, 1971 p.21).

Two short forms of the BTBC (Form A) were compiled and separately scaled for children between 3½ and 5 years. Their brevity enables the tester to hold the attention of a very young child throughout the administration of the test.

Highly significant differences were observed between the mean scaled BTBC scores of three broad socio-economic groups. The mean scores of the lowest SES group remain low throughout the infant school years, those of the middle group hover around the sample mean whilst the mean scores of the high SES group remain well above average. However, there was considerable overlap of scores across the groups. There is no evidence that the social class differences, reflected in the BTBC scores, are greater below 5 years than at 7 years of age.

The mean BTBC scores of children from various types of domiciliary areas were compared. The results suggest that children tend to get higher BTBC scores if they live in a rural locality, which is not notably marked by social disadvantage, than do children of the same social class who are in any other type of area.

Children from middle and low SES groups who attend either rural or urban schools outside the inner city which are in areas marked by social difficulties perform less well on the BTBC than children in localities which are not markedly disadvantaged. Children from urban rehousing estates tend to perform less well than do their counterparts in the inner city or in other urban areas. The performance of children in inner city primary schools was higher than many earlier studies suggest.

Teachers' discourse was observed and curriculum material examined to establish the content validity of the 50 BTBC items for children in England. Though mastery of all the fifty concept terms was not found to be essential for success in British first schools, most children are likely to encounter them at some time before their eighth birthday.

A sample of teachers were asked to assess their pupils' achievements some nine or twelve months after the children had been tested with the BTBC (Form A). Correlation coefficients of BTBC scaled scores and teachers' assessments of mathematics (counting and computation), mathematical concepts, spoken language and reading were, respectively, .481 ($N = 401$), .457 ($N = 371$), .449 ($N = 428$), and .424 ($N = 388$) ($p < .01$ in each case).

After calibration and test of goodness of fit, the Rasch latent trait model was rejected as a suitable method of analysis: the BTBC items are unlikely to be unidimensional. Principal components analyses produced a large number of factors but yielded no clear major factors. The first factor accounted for only a small percentage of the total variance (ranging from 9.0% at 7 years to 15.4% at 5 years of age). Neither Varimax nor Promax rotation produced consistent factor loadings on the test items across the age levels.

The BTBC is considered to be a multidimensional test which estimates children's general cognitive development as well as being a guide to their understanding of the Boehm concepts.

Analysis of each of the 50 BTBC items suggests that although the two alternate forms of the BTBC yield roughly equivalent total scores there are substantial differences in the facility values and in the content of a number of items.

The errors made on each item suggest that non-linguistic preferences and strategies, as well as semantic, syntactical and conceptual knowledge, influence children's responses to the BTBC items. As a result of non-linguistic information and limitations in the design of certain items there is considerable diversity in the efficacy of each item. It is considered that some items do identify the children who understand both the concept and its label and also the children who lack such knowledge but there are other items for which a pass indicates understanding but a failure may not necessarily mean that the child lacks conceptual understanding of that particular concept. With other items, failure indicates lack of knowledge but a pass does not necessarily indicate that the child does understand the concept. With a few items, neither success nor failure provide reliable information about a child's understanding of that particular concept.

Nevertheless, used in conjunction with other information, it was concluded that the BTBC is a useful test instrument.

CHAPTER 2

Statement of the Problem

Investigation of the effectiveness of education, or indeed any research which involves assessing young children's cognitive skills, requires reliable measures of achievement or ability. In the realms of language and conceptual development there is a dearth of measuring instruments which are quick, simple to administer and score, and are standardised with young British children. Linguistic and conceptual understanding can be assessed and analysed by tests such as the ITPA, the WISC, the Stanford-Binet and British Ability Scales but these require skilled testers and are time-consuming. Where large population samples are used, and educational psychologists are not available to administer the tests, a short test of basic concepts could be a useful tool. It could provide an additional instrument for a test battery. The Boehm Test of Basic Concepts, given an English standardisation, might fulfill this need.

The Boehm Test of Basic Concepts (BTBC) was designed to measure children's mastery of a group of concepts considered to be necessary for achievement in the first years of school (Boehm, 1971). Boehm states that the purpose of the test is "to identify

(a) individual <u>children</u> whose overall level of concept mastery is low and and who therefore may need special attention, and

(b) individual <u>concepts</u> with which large numbers of children in a class may be unfamiliar.

In this way the test is intended both as a 'detector' and as an instructional device for use by the classroom teacher" (ibid., p.4).

The BTBC directions are straightforward, the scoring unambiguous and administration requires no special training. The materials consist of two booklets of pictures. The child is required to indicate which of several objects is "in the <u>middle</u>", is "<u>half</u> gone", has "...in <u>every</u> bowl" etc.. There is no need for the child to speak so speech production difficulties are not an obvious handicap to performance on the test. It is quick to administer and the children enjoy it (Dahl, 1973; Smith, 1977).

At present the BTBC lacks standard scores and provides no norms for British children. The manual provides percentile equivalents of raw scores by socio-economic level for the American school grades kindergarten, 1 and 2. The present investigation seeks to establish standard scores for children in England between the ages of 3:6 and 7:6. The aim is to provide standard scores for each half-year group. Whether separate norms for different socio-economic levels are necessary, or feasible, in the United Kingdom remains open until the results have been analysed.

Before a standardisation of the BTBC is put to use a number of questions need to be asked about the test's validity and reliability. Firstly, what precisely does the BTBC measure? It is uncertain whether it measures children's understanding of the concepts involved together with their knowledge of the linguistic labels employed by Boehm or merely the latter. Do perhaps some children understand the concepts but label some of them differently?

Secondly, are the BTBC terms the ones that teachers assume that children understand as Boehm claims? If her claim is correct for American education does the same hold true in Britain? Also, are the fifty BTBC concepts commonly in use in British classrooms? The extent to which the concepts the BTBC is designed to measure are representative of the range of concepts understood by young schoolchildren requires consideration. Bias against any sub-culture should be looked for.

Thirdly, Boehm's (1971) claim that mastery of these concepts is necessary for achievement in the first years of school requires examination. Fourthly, whatever the BTBC measures, its consistency or internal reliability for the British population needs verification.

Standardisation of any test requires a large population sample which is representative of the total population. Responses from a representative sample of nearly two thousand children are likely to yield information about the achievement levels of various socio-economic groups and groups living in differing urban and rural environments.

Evidence suggests that the lower the socio-economic class the lower the verbal achievement as measured by standard tests (e.g. Davie, Butler and Goldstein, 1972; Wedge and Prosser, 1973; The Consortium for Longitudinal Studies, 1983). Therefore it is predicted that the mean BTBC score of each social group will be lower than that of each group ranked higher in the social scale as assessed by parental occupations.

There is evidence that nursery experience can increase the verbal test scores of intellectually disadvantaged[1] children (The Consortium for Longitudinal Studies, 1983; Smith, 1977). So it is predicted that social class differences in BTBC scores will decrease with length of schooling. However, should this turn out to be so, it cannot necessarily be concluded that the disadvantaged are catching up with their more privileged peers. It may mean that the test had underestimated the more disadvantaged children when they were younger. Standardised tests are seen by some writers as having a middle-class bias (Labov, 1969). This applies not only to content but to the comparative ease with which the middle-class child falls into the role of test respondent (Tizard, 1974). Bernstein points out that the middle-class child is more practised at "giving selective attention to an adult" (ibid.). Improved achievement, as measured by tests, may be due to increased self-confidence and motivation allowing children to function at their best level during the test (Zigler and Butterfield, 1968). So whether

[1] Throughout this thesis the term 'disadvantaged' will be used as defined by Tizard (1974) i.e. 'socially disadvantaged' will refer to children from certain social groups e.g. social class IV and V, large families, single parent families, residents in certain areas. 'Intellectually disadvantaged' will refer to children with certain characteristics such as low scores on language and/or other cognitive tests which are likely to result in school failure. 'Educationally disadvantaged' will refer to children who experience learning difficulties at school, leave early and do not enter tertiary education.

any decrease in social class differences, in the BTBC scores, which may be observed in the older children will be because of real cognitive advance on the part of the less privileged children or will be the result of their increased ability to respond to tests will remain conjectural.

It is believed that inner city children are at a particular intellectual disadvantage (Plowden, 1967; Halsey, 1972) though neither social nor intellectual disadvantage is confined to the inner city (Smith, 1975). In Britain since the post-Plowden and Educational Priority Area research in the early 1970s there has been virtually no further evidence published which relates young children's verbal attainment to the type of area in which they live. Therefore it is deemed desirable to compare the BTBC scores of children at schools in inner city and other urban areas, and in socially disadvantaged rural and non-disadvantaged rural areas.

It has been suggested that in early childhood girls are more linguistically mature than boys (Pringle, 1974; Jersild, 1960; Gesell, 1940). However, "the evidence of consistent sex differences in language development is too tenuous and self-contradictory to justify any claims that one sex is superior to the other. In the present state of language assessment the only tenable position is that there is NO significant difference between the sexes in linguistic ability" (Macauley, 1977). By reworking the data of Maccoby and Jacklin's classic study, Plomin and Foch (1981) computed that the sex difference claimed accounts for only 1% of the total variance of verbal ability. These authors, however, find significant sex differences for three cognitive tests (vocabulary, WISC-R Block Design and Picture Memory) but, again, they account for less than 5% of the total variance (ibid.). Plomin and Foch suggest that sex differences may possibly be important at the extreme ends of a distribution. Previous studies report negligible or no sex differences in performance on the Boehm test (Boehm, 1966; Brown, 1975; Ernhart, Spaner and Jordan, 1977; Levin, Henderson, Levin and Hoffer, 1975). Boehm (1971) makes no mention of sex differences in her standardisation sample. Therefore, it is predicted that no significant differences will be found between the sexes in their scores on the BTBC.

There is evidence that infant school teachers rate girls higher than boys in speech development (Brandis and Bernstein, 1974; Hartley, 1980). Such expectations may influence achievement and test performance. Thus the research data will be analysed in order to detect any sex differences which may occur at particular ages.

Data from such a large sample should provide some insight into young children's non-linguistic strategies when faced with a difficult task. It may be that, when uncertain of the correct response, children who have learnt the left to right orientation required in reading will tend to choose the first in a row of pictures. Alternatively, young children's responses may be influenced by their preferences for the larger amount or the topmost surface as suggested by Eve Clark (1973, 1977, 1980); or they may simply choose the object they would most like to have such as an ice-cream or a doll. The wrong responses, as well as the correct ones, will be examined for evidence of such strategies.

By analysing the children's responses to the BTBC items support may be found for Clark's (1973) semantic features hypothesis. This hypothesis assumes that in the early stages the child's use and interpretation of words may differ considerably from the adult's. Clark posits a fixed order of acquisition of word meanings starting with the most general features and moving towards more specific ones. For example, 'big/little' is acquired before 'tall/short' and so on. The positive, linguistically unmarked, dimensional and locative terms (e.g. more, high) are acquired before their negative, marked opposites (e.g. less, low). "It would not be until the child learnt the contrasting values of polarity - that more was positive and less negative - that this pair of antonyms would be interpreted correctly" (Clark, 1973, p. 73). Prior to this the meanings would be confused. Thus, according to the semantic features hypothesis, BTBC items such as '... the most eggs' should be easier than items such as '... the least stars', especially for the younger children.

In short, this investigation is concerned with three domains viz.

1. The standardisation of the BTBC on children in England and the examination of its reliability as a measuring instrument, together with its validity in assessing children's concept acquisition.
2. The establishment of relationships between success on the BTBC and the child's sex, socio-economic group and type of locational area of the school he attends.
3. Children's non-linguistic strategies as revealed by their incorrect responses as well as the order of acquisition of relational terms as revealed by their correct responses.

CHAPTER 3

Review of Literature

(i) Conceptual Development in Early Childhood

Introduction

This chapter is divided into two sections. In Section A theories of cognitive development in young children are reviewed. Obviously a comprehensive review of all the literature on this vast subject is not possible. So a selection has been made from the main body of theory which informs our understanding of how young children form concepts.

No attempt is made to encompass all of Piaget's elaborations, for example, nor all of the prolific criticisms and counter-criticisms of his opponents and supporters. The review concentrates on Piaget's more recent writings rather than on his earlier work. Some studies which suggest revision or expansion of Piagetian theory in relation to early conceptual development are considered.

Social learning and behavioural theories are briefly reviewed and evidence, from the Follow Through project, of the efficacy of behaviourism in the classroom is evaluated. Reference is made to both Gagne's and Vygotsky's delineation of stages in concept formation and some of the more recent theories of concept development are briefly described.

Fischer's (1980) exposition of skill theory is reviewed as an exemplar of this body of research. Klausmeier's (1979) Conceptual Learning and Development (CLD) theory is reviewed as an important model which combines Piagetian, skill theory and information processing (IP) constructs. Finally, some models of cognitive development and learning produced by the information processing approach are reviewed together with some of the criticisms made of the formal models and of IP methodology.

In Section B psycholinguistic theories and studies of young children's language development are reviewed.

Though 'language development' and 'cognitive development' are closely related it became necessary, for the sake of literary organisation, to make a division. Theories concerning the relationship between language and cognition are reviewed. This inevitably leads to consideration of concept organisation and it could justifiably be argued that much of the material in Section A belongs here and vice versa. To avoid repetition and undue wordiness the connections and overlap are left implicit.

Theories about young children's semantic development are then considered and, in particular, Eve Clark's semantic features hypothesis is described together with a number of studies which it has generated. Syntactic and phonological theory are deemed only marginally important to the theoretical basis of the BTBC so the early syntactic studies of Roger Brown and Lois Bloom, for example, are omitted. Nevertheless some consideration of syntactics together with pragmatics is included.

Section B ends with a selective review of sociolinguistic theory in the wake of Bernstein's early writings or, at least, that part of the theory which has been employed in the 'cultural deficit' versus 'cultural difference' debate.

Section A. Cognitive Development in Young Children

Piaget's research is centred on the relationship between 'the knower' and 'the known'. The subject's encounters with the object (thing or person) are investigated in order to describe and then explain the structures which underlie the child's behaviour and determine what he can do (Vuyk, 1981a).

Piaget reexamined earlier models, which he found insufficient, and in 'The Development of Thought' (1977a) puts forward a causal model for mechanisms or functions of development.

"Knowledge proceeds neither solely from the experience of objects nor from an innate programming performed in the subject but from successive constructions, the result of the constant development of new structures." (Piaget, 1977a p.v.).

This is the central idea on which Piaget based his theory of equilibration.

Assimilation and accommodation are still seen as the two central processes which form the components of cognitive equilibrium. Equilibriums are seen as stationary yet dynamic states. Equilibrium does not "constitute a stopping point, since any finished structure can always give rise to new requirements in fresh substructures or to integrations in greater structures (ibid., p.30).

"To conceive of equilibration as a mere step to equilibrium would ... be inadequate since in addition to being a means of structuring, it is constantly attempting to achieve a <u>better</u> equilibrium." (ibid., p.31). Piaget now speaks of "increasing equilibriums". "Nonbalance plays a functional role of prime importance since it necessitates reequilibrations." (ibid., p.178). Affirmations are compensated by the construction of negations resulting in alternating periods of equilibrium and disequilibrium (i.e. the process of equilibration). The ensuing period of equilibrium is at a higher level (equilibration majorante). It is seen as 'higher' because (a) a structure with a greater number of substructures is more stable and (b) because the subject's interaction with the environment extends.

Piaget distinguishes between three forms of equilibrium: that of relations between subjects and objects, that of coordinations between schemes or subschemes and the general equilibrium between the whole and their parts hence their integration into a total system. He suggests that it is this third equilibrium, dominating the other two, which appears to determine action finality. New forms deal with previous ones and include them as contents. "The process of this endogenous construction uses reflecting abstractions which take the elements of new forms from the heart of the most elementary ones" (ibid., p.179). He elaborates further:

"The improvement of the equilibration then is possible because the superior system is the seat of new regulations and its construction includes a more complex set of assimilations and accommodations... The later schemes are richer than the preceding, since the reflecting abstraction leads to a greater number of compositions and this increased wealth of regulations leads to improved control. The result is a hierarchy of regulations of regulations leading ... to self-regulation and self-organisation" (p.179/180).

Piaget seeks an explanation of why these new constructions occur. From the statics branch of physics he borrows the principle of 'virtual work' and shows the equilibrium of the cognitive system as consisting of a compensation of all the 'virtual work' which is "compatible with all the modifications possible under the constituent laws of the structure" (ibid., p.180). In this general equilibrium of the whole and its parts the disturbance is only related to 'virtual work'. Each completion of a construction opens new possibilities which did not exist as such at the previous levels (ibid.).

The next step in the explanation suggests that if a new possibility imposes itself as the necessity to go beyond compensating a virtual non-balance it did not "arise by chance from the subject's spontaneous inventions, but (it) is a product of the multiplication of the independent subsystems..." (ibid., p.182). The increasing coherence of the subsystems conditions this "third kind of equilibrium which imposes itself between the general integration and the differentiations" (ibid., p.183). These subsystems, which at first develop relatively independently, interact in a complex manner resulting in further elaboration of the structural features common to the substructures in question (Inhelder et al., 1974).

Thus the reconstruction process means far more than the addition of a greater number of elements of the lower type in order to reach the higher level (Vuyk, 1981a).

The learning studies of Inhelder, Sinclair and Bovet (1974) were designed to gain further insight into the transition, or construction, mechanisms which enable children to attain conservation and class inclusion concepts, in short, the acquisition processes. No attempt was made to lead the child through a series of preprogrammed steps toward the correct solution to a problem. Instead, facilitating situations were provided. These led to comparisons and conflicts between the subjects' predictions and ideas and the actual outcome of manipulations. Study of the results provided evidence that "new acquisitions are not the result of a simple generalisation of previously acquired knowledge to a new context, but of a true reconstruction on a new level" (ibid., p.253).

This involves a transposition of conflicts and strategies and was seen to take place in four steps. At first the child kept two modes of reasoning apart (e.g. spatial ordinal correspondence and numerical correspondence) i.e. separation between schemes. In the second phase the child has an apparent urge to understand the discrepancy of his solutions i.e. juxtaposition of schemes. Then he tries to reconcile them and compromise solutions are used in the third stage. Such solutions may be said to be half-way between preoperatory and operatory reasoning. The solutions are obtained through incomplete, or partial, compensation. The reciprocal accommodation and assimilation are achieved only during the fourth phase. Now there is complete compensation (ibid.).

The incomplete compensations, or compromise solutions, demonstrate the child's own activity in the construction of new forms of knowledge. Thus "the situations most likely to elicit progress are those where the subject is encouraged to compare modes of reasoning which vary considerably both in nature and complexity, but which all, individually are already familiar to him. Training procedures in which one type of reasoning is artificially isolated and exercised, as is often the case in certain programmed learning projects, are not, in our opinion very useful, since they eliminate the element we are considering necessary for progress i.e. the dynamics of the conflict between schemes...

"Not only are disturbances and imbalances mentally compensated for, but new constructs are established through regulatory mechanisms, which themselves undergo further development during the acquisition process" (ibid., p.265).

The results of Inhelder's training experiments were in accord with Piaget's model: the same material only led to an experienced conflict when the child had already reached a favourable level of development. Otherwise there was no perturbation and no progress (Vuyk, 1981a).

"Concepts do not develop in closed systems but are in constant interaction with each other and it is this interaction that accounts for the child's progress during learning experiences." (Inhelder, 1974 p.15). Piaget holds that concepts, being more complex than perceptual data, presuppose operations that are abstracted from actions and their coordinations. These operational concepts are general and reversible. As a child can act correctly before he is able to verbalise what he has just done the process of concept acquisition progresses from action to conceptualisation (Vuyk, 1981a).

The essence of conceptualisation, or cognizance, is examined by Piaget in "The Grasp of Consciousness" (1977b). Piaget distinguishes an "elementary consciousness" from a "cognizance". Cognizance constitutes the passage from the action to its representation and includes a reconstruction arising from the conceptualisation. Elementary consciousness is linked to an immediate observation, is fleeting and gives rise to no conceptual or representative integration. The semiotic function enables a perception to be accompanied by an interpretation. This interpretation (i.e. any form of verbal or imagined conceptualisation) enables the perception to be integrated and that constitutes its cognizance (ibid.).

The mechanism that renders conscious the hitherto unconscious elements consists of "a passage from practical assimilation (assimilation of the object into a scheme) to an assimilation through concepts" (ibid., p.337). In this process the apparent degrees of consciousness

"depend on different degrees of integration. For example, what precedes conception ('subception') and is defined as 'unconscious perception' could well be accompanied by a certain consciousness at the moment it actually occurs, but this remains temporary in the sense that it does not seem to be integrated into the subsequent states" (ibid., p.342).

"In sum, the mechanism of cognizance appears ... as a process of conceptualisation - reconstructing, then going beyond, on the semiotic and representational planes, what was acquired on that of the action schemes" (ibid., p.342).

Piaget's hypothesis that action is an autonomous form of knowledge which only becomes conscious later was generally confirmed by the experiments described in "Success and Understanding" (Piaget, 1978). Here, the child manipulates the material in order to reach a goal set by the experimenter. He is then questioned about the reasons for his actions. The results led Piaget to impose a restriction on his earlier hypothesis about action: at a higher level (about 7 to 8 years onwards) conceptualisations influence actions. Succeeding in an action, and far more, failing to do so, stimulates the subject to ask 'why?'. The answer gives him reasons which constitute understanding and thus allows him to anticipate future actions and to plan (ibid.).

"If this work were to be followed up from outside the Genevan framework it would be important to establish what happens when children are allowed to formulate explicit hypotheses spontaneously rather than under questioning from the experimenter, and also to begin to identify what input from children's physical or social environment is relevant to changes in the relationship between knowing and doing" (Robinson, 1983 p.131).

Any reading of Piaget begs the question: what is action? In his earliest writings 'action' denoted fine motor activities. Later he included physical activity such as eye and head movements. 'Action' became defined as 'purposive movement'. Piaget also said that the repetition of a scheme which gives the child nothing new is not an action. In his work on correspondence he stresses that even without a transformation there is still 'action' i.e. the act of comparing. Discussing affirmations, Piaget writes:

"Now it is clear that, though an action is always directed towards an aim... and though this aim is essentially positive ... it is nevertheless a fact that every action tends to change or negate an initial state in favour of a new one ." (Piaget, 1978 p.228).

Beilin argues that it is impossible to distinguish between processes that involve action and those that do not as all psychological processes can be said to involve action (Vuyk, 1981b). Vuyk suggests that the emphasis should not be on 'action' but on the constructive use of action.

From a review of literature Anthony (1977) concluded that Piaget's insistence on physical activity and dismissal of observation are excessive. Anthony cites Flavell, Hooper, Beard, Ginsberg and Opper and Piaget himself in support of his contention that the main practical implication drawn from Piaget's theories is the insistence on physical activity for intellectual development. He examined six training experiments in teaching concrete operational thinking without involving physical manipulation by the child. Each resulted in some transfer of training though none fulfilled all the conditions laid down by Inhelder et al. (1974) for evaluating Piagetian training. None of the experiments was perfect but Anthony is satisfied that together they indicate support for his conclusion that the extreme Piagetian insistence on physical activity has been excessive. He writes:

"the credibility of his theory is somewhat diminished when a hypothetical assertion derived from it by its author appears to be false, and especially when the hypothesis seems to be the theory's main practical implication." (Anthony, 1977 p.22).

The crux of the matter seems to be the child's involvement (Inhelder et al., 1974). This is not explored by Anthony who sees a dichotomy between Piaget's statements on the necessity for action and Inhelder's on children observing other people's actions (op cit). Involvement is necessary if the child is to experience discrepancies between predictions and outcomes. The necessity of such involvement is reflected in publications which have had a considerable influence on current Nursery and Primary practices in this country (for example Plowden, 1967; Brearly et al., 1969;

Tough, 1973; Parry and Archer, 1974, 1975; Manning and Sharpe, 1977; the Nuffield Mathematics Project; Science 5/13, 1972). These authors, and many others, talk about the child's active involvement, active exploration and active observation. A random search of such literature does not reveal that child-centred educationists devalue observation. Like Piaget and Inhelder (op. cit.) they distinguish between 'passive' and 'active' observation or action. Piaget's epistemology has always posited the constructive interaction of the child and his environment (Karmiloff-Smith, 1978).

Nor have leading Primary School practitioners and trainers interpreted Piaget in the way that Anthony suggests. Rather they have used Piaget's investigations to support their stance already arrived at by philosophical theory and intuitive judgement based on their own experience (Brearley and Hitchfield, 1966).

The many practical implications of Piaget's work are too numerous to describe here but they have been acted on to improve the quality of adult intervention in young children's learning, to increase the challenge and stimulation of the environment provided, to support the need for a flexible and individual approach to each child's learning based on acute observation by the teacher. The latter, to a large extent, derives from Piaget's clinical method (Plowden, 1967; Brearley et al., 1969; Tough, 1973; Parry and Archer, 1975; Poulton and James, 1975; Manning and Sharpe, 1977). The emphasis is on first-hand experience and a recognition of the limited value of rote learning (Plowden), on the individual child as learner and on his thought processes rather than on the end results of 'programmed' learning (Brearley et al., 1969). It is observed that children may fail because they cannot actually use concepts that they have memorised (Vuyk, 1981a; D.E.S., 1978).

Six general implications of Piagetian theory for educational practice may be summarised as follows:

"1. The language and thought of the child are qualitatively different from that of the adult, and the child may find it difficult to assimilate some sequences of ideas.

2. The young child learns more easily from concrete experiences.
3. Tasks should take into account the child's readiness for, and interest in learning.
4. The development of curricula must take into account the invariant sequence of cognitive development.
5. Social interaction is a crucial factor in the child's development.
6. Traditional methods of instruction - that is, group lessons with a given sequence of material, transmission of materials via lecture or other verbal explanation - have grave deficiences."

(Lawton and Hooper, 1978 p.175 after Ginsburg and Opper).

These basic principles "are perhaps best exemplified by Kamii (1973) who succinctly states that learning is an active process, that social interaction is important for intellectual development, and that intellectual activity is primarily based on actual experiences rather than on language per se" (ibid., p.175).

More recently, Kamii and De Vries have developed the Piaget for Early Education Program which, unlike most American programs, does not provide a 'cookbook' curriculum but lists guiding principles such as 'encourage children to be independent and curious', 'encourage children to interact with their peers' (Brainerd, 1983). Unfortunately no evaluation data is available (ibid.).

Piaget's theory has been with us for some sixty years despite almost continuous criticism but, as Beilin (1980) points out, it appears in no immediate danger of being superceded and "it is not clear what theory would substitute for the large domain of cognitive phenomena that Piaget and others have described" (ibid., p.246).

Beilin argues that although "functionalist approaches such as those of Bruner, Olver, Greenfield and others (1966) posit theory like proposals for restricted aspects of development, it is questionable as to whether they represent a theory of development as such" (ibid., p.246). He asserts that much of the essence of functionalism and of information processing theory has been or could be incorporated into Piagetian theory.

Boden (1982) discussing the importance of Piaget's theory of equilibration attests that despite its vagueness it

"identifies an important theoretical problem - how the generation of harmonious novelties is possible - which is still highly obscure, ... He was mistaken in thinking that as well as indicating the problem he had provided its solution. Understandably, many critics regard it as well-nigh devilish on his part to have tempted his disciples into the illusory faith that 'equilibration' provides an explanation. But to point out significant problems that are largely unrecognised by other theorists is no mean feat" (ibid., p.172).

Piaget's theory has not remained static. Revisions and expansions were made by Piaget himself to accommodate new experiments and data. (See Piaget, 1977a, b, 1978; Beilin, 1980) Further refinements have been and are currently emanating from Geneva. (Those concerned with language are reviewed below in Section B.) Some of the challenges to orthodox Piagetian theory which are particularly relevant to young children's conceptual development will now be considered.

Meadows (1983) notes the lack of empirical evidence for 'structures-d' ensemble':

"though there are an enormous number of replications of most of Piaget's sets of data on such things as conservation and classifications there are very few indications of how they fit together. It seems unlikely that there really are 'structures-d'ensemble'.

It is still, of course, possible to describe thought in terms of logical operations, even if the operations do not fit as tightly as Piaget sometimes implied..." (ibid., p.p.18/19).

Gelman (1982; 1979) is among those who challenge some of the Piagetian claims about the egocentricity and perspective - taking abilities of pre-school children, their number and counting prowess, their classification capacities, their sensitivity to temporal order and their ability to use rules to solve problems. She writes

"...once one begins to talk about precursors of later cognitive abilities it is no longer unreasonable to start the search for those concepts and capacities the preschooler must have if he or she is to acquire complex cognitive abilities. We should expect to find domains in which they are quite competent - if only we look." (Gelman, 1979 p.902).

Gelman (1982) suggests that access theories of development "offer a potential setting within which to capture the picture of early cognitive development..." (p.218) but preschool children are at a disadvantage:

"Theirs is a competence that is fragile, that can be on-again, off-again, that is used only in restricted settings, that does not generalise readily. In other words they have limited access to their competencies. ...The theory leads one to expect pockets of competence which only reveal themselves under 'appropriate' circumstances." (ibid., p.218).

Thus, we "must cease to approach young children with only those tasks that are designed for older children" (Gelman, 1979 p.904). In order to make a full and accurate assessment of present cognitive functioning tasks should relate to the child's age and developmental stage in thinking (Lunt, 1983).

McGarrigle and Donaldson's 'naughty Teddy Bear' experiments and other studies of the effects of situational contexts may elicit evidence of conservation in young children (Donaldson, 1982, 1978; Light, Buckingham and Robbins, 1979; Hargreaves, Molloy and Pratt, 1982; Miller, 1982; Meadows, 1983; see also Chapter 4). Albeit, though preoperational children may have competencies as yet largely untapped, the evidence to date poses no serious threat to the essence of Piaget's preoperational, transitional and operational stage theory.

Numerous training studies have demonstrated powerful training effects on operational acquisitions by young preoperational children (e.g. Bryant, 1974; Vuyk, 1981b). Consequently, Beilin (1980) considers the possibility that Piaget's theory is incorrect with respect to assimilation but regards this as a premature conclusion. He argues that "it would have to be shown first that training studies in principle provide a decisive test of the theory" and "the data have to be conclusive" (op. cit., p.254). Beilin accepts as conclusive the data from many studies but

"There are a much larger number of studies, however, which suggest...

...that:

1. The differential effects of age on training are considerable, that is, older and more advanced operational subjects show superior training effects to younger and transitionally - operational subjects who, in turn, are superior to preoperational subjects.
2. In absolute terms, a number of studies show no training effects with younger preoperational subjects, even when they do show training effects with older subjects (e.g. Bruner et al., 1966)."

(ibid., p.p.254/5).

Furthermore, if one distinguishes between natural and experimental or 'forced' development, Beilin posits that in natural conditions the assimilation constraints exercised by existing schemes will be consistent with Genevan findings. Likewise, training studies

"that are based on nonforced equilibration methods will generally operate under the same constraints. 'Brute' or forced methods, such as those used in verbal rule instruction...will force assimilation to the simplest available structures, but these will have only limited generalization value with the youngest children. However, the more developed preoperational schemes of older, and transitionally-operational children, will provide a more stable base for the assimilation of forced strategies or solution algorithms, or lead to the forced restructuring of incomplete or unintegrated structures. In these circumstances, the training data...need not, necessarily, do violence to Piaget's theory. Instead the theory can be modified to accommodate the research evidence...(ibid., p.p.255/6).

Field (1981) agrees with Beilin. She reports a training study of preschool non-conservers which included a long-term follow-up. Understanding was retained by most of the children but the differences between the three and the four year olds seem to bear out Beilin's observations. Field concludes that games and verbal rules do not lead children "who do not have any constituent structures to become operational. Instead, the training provides a strategy. This strategy is 'affixed to' the emerging structures. Perhaps because it is a useful strategy in life outside the experimental situation, it is maintained by those children whose structures are sufficiently well developed." (Field, 1981 p.333).

Beilin does not dismiss the non-Genevan training studies as having no bearing on the understanding of cognitive change:

"If the effect of training is such that many different types of experience can lead to the same result (for example, conservation) then... no specific element in training or experience can be said to be both a necessary and sufficient source of change that would lead to conservation (or other cognitive operations). ...what is assimilated is essentially 'content' (e.g. information about the physical properties of objects or events); structure, however, is internally constructed." (Beilin, 1980 p.p.256/7)

Piaget's equilibration theory posits conflict as the spur to intellectual development (Piaget , 1978). Bryant (1982) tested the hypothesis that it is agreement, and not conflict, between strategies which produces intellectual change. In each of four experiments ten six year olds were either matched or yoked with ten controls. The criterion was whether different treatments provoked more measuring of vertical wooden blocks (one on the floor and one on the table) with a rod which was lying nearby or whether the children continued to rely on direct comparisons. "The main question ... was how a child establishes for himself when and where it is right to use particular strategies. The results support the idea that he does this by checking one strategy against another. If they agree his confidence in them increases: if they conflict he is no further forward." (ibid., p.249) The children measured more after seeing that their measurements agreed with direct comparisons (i.e. placing the blocks side by side). Neither experience of conflict nor telling children that their measurements were right had any effect on their willingness to measure.

Bryant admits the limitations of this evidence: the numbers of children were small, the experiments were confined to one age group and one narrow cognitive skill and there were no tests of permanency. However, controls were stringent. One wonders if the same results would be forthcoming in real life. Was the experimental situation sufficiently motivating for six year olds to produce real perturbation? Generally, direct comparison is more accurate than measuring and serves the purposes of most six year olds in their constructive activities. When this is impossible, as in buying the right length of 'Contact' to cover a shelf, the consequent perturbation produces the need for measurement. (See Chapter 4 for further discussion of situational contexts.)

Bryant argues that there is only weak evidence for the theory that conflict lies at the heart of intellectual development (ibid.) but, so far, he has not produced convincing evidence for a contrary viewpoint. Moreover, however convincing the evidence, it would be dangerous to generalise from experiments with one task to strategies applicable to other tasks. (See Breslow, 1981)

A refinement of the functions of assimilation and accommodation in the process of equilibration is convincingly argued by Block (1982). He suggests that we are

"programmed to follow the adaptive imperative: 'Assimilate if you can; accommodate if you must!'... It is only when assimilation fails, for whatever reason, to reconcile intrusions within the structured status quo that the destructured individual attempts the second line of adaptation, accommodation. Whereupon the formulation of new structures, if effective, can be used to assimilate what before was unassimilable, thus again reequilibrating the individual... The motivational functions of both assimilation and accommodation are to advance or to restore psychological equilibrium .. Piaget's asserted but unused idea of an equilibrium between assimilation and accommodation has been set aside. Instead, it is posited (as Piaget sometimes suggested) that the assimilative mode has a built-in priority over the accommodative mode..." (ibid., p.286).

Block offers an interesting extension of these principles from cognitive theory to personality functioning:

"...we may distinguish individuals with respect to their readiness or reluctance to switch from the assimilative mode of processing experience to the accommodative mode when disequilibration arises. ... The ordering of individuals with regard to their degree of continuation with insufficient assimilative structures before attempting accommodation represents, within the present perspective, a primary conceptual variable for the study of individual differences." (ibid., p.292)

Block discusses some implications such as the connections between socialisation practices and the anxiety-inducing efforts of innovative accommodation and makes a plea for more attention to individual differences.

Experiments by Emler and Valiant (1982) contribute to both the conflict versus agreement argument (Bryant, 1982) and to the need for empirical psychology to pay more attention to individual differences (Block, op. cit.). Social interaction and cognitive conflict in the development of spatial coordination skills were examined. The results supported the proposition that social interaction stimulates cognitive development in so far as it confronts the child with the experiences of disagreement.

Nevertheless, opposing points of view do not always produce disagreement and even when disagreement is manifested it does not necessarily produce cognitive conflict or developmental change (ibid.). Emler and Valiant consider possible reasons: 'egocentric' communication failure where children blame the listener for failing to understand or where they have difficulty appreciating that they rather than the other may be wrong or that both may be wrong.

Additionally, they speculate on "another set of factors that have been relatively ignored in this area of research, those of the relationship that exists between the children who are required to interact" (ibid., p.302). Existing relationships of friendship, enmity, or dominance are seen as influencing factors:

"It may also be that children are more surprised when they find themselves in disagreement with friends than with others and for this reason, or perhaps reasons associated with the greater facility of interacting with friends, will profit more from such disagreements. Rather than ask whether social interaction as such is important, we should perhaps in future ask: interaction with whom and disagreement with whom contributes to development?" (ibid., p.302)

A series of experiments by Doise and his colleagues with children from about 5 to 9 years of age established that the performance of children when interacting with peers was typically at a higher level than when they were given solitary practice on a task (Light, 1983). Furthermore, Doise et al. claim that more is involved than simple imitation; "the Piagetian concept of conflict between individual centrations is used to account for progress" (ibid., p.72). "It does not matter if both partners are equally wrong, as long as they are wrong in different ways" (ibid., p.72).

Light suggests that conflict is important "not for the reasons that Piaget and Doise have supposed. Rather the conflict serves to shift the issue from the subjective to the objective ... The crucial function of social interaction ... is to establish or clarify for the participants the nature of the questions that are being asked" (ibid., p.76).

After examining a number of studies of social interaction and cognitive development, Light sees the picture as "one of convergence of findings at the empirical level but divergence of views at the theoretical level" (ibid., p.85). Ongoing and future research is likely to lead to a rapid development in understanding of these issues (ibid.).

The role of other people in children's concept formation has long been emphasised by social learning theorists (e.g. Bandura, 1967, 1977) from a very different theoretical base. Children are thought to acquire concepts through both social feedback processes and the observation of models. Concepts, once abstracted, subsequently influence the child's acquisition and interpretation of future environmental information and the acquired concepts are in turn refined by subsequent social experiences (Bandura, 1977). Modelling procedures, it is claimed, are effective in training children in problem solving strategies and even more successful when the model verbalises the rule being used (Brody and Zimmerman, 1978). Brody and Zimmerman claim that when teachers were taught to use modelling and corrective feedback, to teach the concept of seriation, instead of their customary "Socratic type of questioning technique" they were significantly more effective and their pupils "exhibited acquisition and transfer of the concept on novel stimuli" (ibid., p.471).

Social transmission theory "regards the laws of logic as a set of cultural conventions which are passed on to the child via the linguistic system as part of the process of enculturization" (Russell, 1982, p.265). Discussing the role of linguistic processes in the growth of necessity, Russell comments

"The main difficulty with this approach is that it has few conceptual resources for explaining why the child comes to regard certain conventions as necessary vis-à-vis non-logical conventions such as those of a legal system or of phatic speech" (ibid., p.265).

Corrective feedback is also a fundamental principle of Behaviour Modification techniques as exemplified by the Behavior Analysis and Direct Instruction models in Head Start and Follow Through programs. The outcomes of nine Follow Through programs are analysed by Becker (1978). Five of these were based on individualised approaches which focused on the whole person and encouraged child-initiated activities, two were behavioural programs: the Behavioral Analysis Model which focused on the systematic and precise use of positive reinforcement and the Direct Instructional Model (DISTAR). The remaining two programs shared common features with both these major groupings; one was the Language Development (Bilingual) Model (SEDL) and the other taught parents to teach their children. Becker concludes that

"The various data analyses comparing different approaches to teaching the disadvantaged show rather overwhelmingly that the Direct Instruction Model was more effective in term of basic academic goals and affective outcomes. The Behavior Analysis Model was especially effective on Math Computation and Affective Measures in terms of ISOs[1] and on Reading and Spelling in terms of normative performance levels. The Parent Education and SEDL programs were the only others to show a few positive outcomes." (ibid., p.p.451/2)

However, Becker draws attention to implementation problems with some programs and until this issue "is faced squarely in the design of evaluation studies, conclusions about what works will have to continue to be qualified" (ibid., p.452). There was neither evaluation of model implementation at each site nor of control groups (ibid.).

One problem in assessing the effectiveness of particular teaching models is that of knowing how far the individual classrooms concerned are exemplars of that model (Rosenshine and Furst, 1973). One of the Follow Through models is described as the Open Education Model. "This model is derived from the British Infant School revolution" (Becker, 1978 p.433). In six nurseries, which could each be thus described, the behaviour of teachers differed widely with significant differences in affective outcomes and non-significant differences between the nurseries in children's language and conceptual development (Smith, 1977). It was also found that intellectually disadvantaged children, as defined by the ITPA, could enter

[1] ISO: an index of positive and negative effects of a program called Index of Significant Outcomes (see Becker, p.439).

the infant school at less of a disadvantage in psycholinguistic skills than at the start of their nursery year without the use of special intervention programmes. One might suppose that the DISTAR and Behavior Analysis models with their tightly programmed schedules were less dependent on the calibre of the teachers than were the Open Education, Weikart and other flexible models where success is likely to be more dependent on the insight, imagination and sensitivity of the teachers. Without considerable knowledge of the quality of the teaching in these classrooms it is difficult to comment on their relatively poor showing in the Follow Through evaluation. As Becker admits, "some of the poor outcomes may not be model program effects, but due to implementation problems"(op. cit., p.452). Stringent criticism of the Follow Through evaluation is made by House, Glass, McLean and Walker (1978). Much of their criticism is challenged by Wisler, Burns and Iwamoto (1978) but these authors acknowledge that the "identification of successful sites, combined with the often weak or variable model effects, suggests that local conditions, such as children whose needs match especially well what the model can provide, local variations of the model, or especially skilled teachers, were more apt to determine success than the models used" (ibid., p.178).

Becker's breakdown of "the Basic Skills ISOs by variables ... shows that the positive ISOs for Direct Instruction were largely achieved through successful outcomes for MAT Language (+ 656) and MAT Computation (+ 562)" (op. cit., p.441). That disadvantaged DISTAR pupils scored relatively highly on the Metropolitan Achievement Test is not surprising considering how well they would have been drilled in question and answer techniques in these areas of language and number. (See Leinhardt and Seewald, below) Moreover, Becker and Engelmann organised testing for "every child in the Direct Instruction program at entry, from 1969 to 1973, and each spring thereafter through third grade. ...all children were tested with the Wide Range Achievement Test (WRAT) and the Slossen Intelligence Test (SIT)...

Beginning in 1972, the (MAT) was used at the end of first, second and third grades" (Becker, 1978 p.p.447/8). So the DISTAR children were also practised in test taking. How often, or if, children in other programs received extra testing is not revealed in Becker's article. Moreover, the evaluators' use of MAT scores is criticised:

"They present test scores as if these were adequate measures of generalized constructs; cognitive development, for example, is equated with scores on certain subtests of the MAT. These subtests do not come close to representing the richness conveyed by the concept of cognitive development. The evaluators appear to have a total lack of trust in any data except achievement-test scores even though such child-outcome data present serious problems of interpretation." (Hodges, 1978, p.187).

One such problem is that "an outcome measure may be biased in favor of a particular practice or program because the overlap between the test and one program is greater than the overlap for the other(s). ... When overlap is not considered, one program or set of practices may look artificially better with respect to test performance than another" (Leinhardt and Seewald, 1981).

At first sight the success of behavioural classrooms in terms of affective outcomes is surprising until one considers the measures used. These were the Coopersmith Self-Esteem Inventory (1967) and the Intellectual Achievement Responsibility Scale (IARS + and IARS -, 1965). These have been attacked as unreliable and invalid (Becker, 1978; House, Glass, McLean and Walker, 1978) but Becker retains these test results in his analysis (op. cit.).

Using self-report measures with young disadvantaged children seems highly unreliable. One wonders by how much the children's linguistic abilities confounded the results. A small study of six nurseries, in which the criteria of affective outcomes were increases and decreases in various behaviours of the children over a year, showed that the use of direct teaching techniques such as drill and review correlated significantly with large increases in thumb-sucking whilst the minimum of direction correlated significantly with increases in children's questioning and the sharing of toys (Smith, 1977).

Becker's (op. cit.) claim that follow-up studies at fifth and sixth grade show significantly better performance by DISTAR children than by comparison groups is not supported by Miller and Bizzell (1983). Over 60% of the children who were in four preschool programs for one year were compared at sixth, seventh and eighth grades on IQ and school achievement. "IQs did not differ significantly among the program groups but in reading and math there were differential effects related to both preschool program and sex in all 3 middle-school years" (Miller and Bizzell, 1983 p.727). The four programs were Montessori, Traditional, Bereiter-Engelmann and DARCEE. Fourteen classes were involved with an additional 34 control children.

"At the end of prekindergarten, there was a greater impact on IQ for the two didactic programs (B-E and DARCEE). Thereafter there were no statistically significant program differences, but by the end of second grade, Montessori and controls were the highest groups. Over the 4-year period, IQs declined more sharply for B-E children than for the other program groups" (ibid., p.728).

"There were no significant differences in IQ at either seventh (WISC-R) or eighth (Stanford-Binet) grade for either sex. However, Montessori males were the highest on IQ, while DARCEE males and B-E females were the lowest at both points" (ibid., p.735).

Males who had been in nondidactic programs were superior to didactic males in sixth, seventh and eighth grade reading and mathematics. Females showed no significant differences between the didactic and nondidactic groups except in eighth grade reading which demonstrated didactic superiority.

"It is somewhat astonishing to find that sixth-grade children who were randomly assigned to a particular program at age 4, who were not brighter or more advantaged to begin with, and who had been scattered throughout a large school system during the first 6 years of school were performing significantly better at middle-school than their peers who had experienced other types of prekindergarten programs" (ibid., p.739).

Overall, the girls were found to be performing slightly better than boys except for those boys who had experienced Montessori programs. Miller and Bizzell speculate on the implications of the higher achievements of boys from both nondidactic programs compared with B-E and DARCEE boys. "One possibility is that the behavioral results from didactic instruction serve to reinforce certain tendencies already present in disadvantaged children. For example, it has been shown that lower-class children tend to develop rigid patterns of behavior which may be efficient in the short run but are potentially disabling when task parameters change (Bresnaham and Shapiro, 1972)" (ibid., p.740). Then the authors use the unproven assumption, that boys are less mature than girls at 4, to suggest the possibility that "they were more susceptible to these potential effects of the group instruction offered in B-E and DARCEE and less ready to profit in a lasting way from the excellent skill training offered in those programs. It is also possible that the boys simply needed the more individualized and slower-paced instruction given in the nondidactic programs." (ibid., p.740).

Another comparison of the long-term effects of five different preschool programs (Traditional, Community-Integrated, Montessori, Direct-Verbal directed by Bereiter and Engelmann and Ameliorative directed by Karnes) does not find that any one program demonstrates superior long-term impact over the other four (Karnes, Shwedel and Williams, 1983). Teacher behaviour is suggested as a key variable (ibid.).

Summarising the findings of the Consortium for Longitudinal Studies, in which all the major preschool curricula were represented, Lazar (1983) concludes that

"any well-designed, professionally supervised program to stimulate and socialize infants and young children from poor minority families will be efficacious. It is certainly possible that finer-grained outcome measures might reveal differential effectiveness of the programs, settings, durations, and intensities" (Lazar, 1983 p.462).

The conclusion of most of the Consortium's researchers is that

"chances of later success are raised by preschool but that how much, for whom, how long, and on what measure is dependent, in ways yet to be specified, on the characteristics of the preschool program, the child's circumstances, community support and opportunities, and later school experiences" (Datta, 1983 p.474).

A decade ago it was posited that didactic programmes produce only short term gains and that other methods have more lasting effects (Beller, 1973; Elkind, 1973). Elkind (1973) argued that, if we 'teach' didactically at the preschool level, children may learn by different means than they would do later because they lack the logical and linguistic structures that they would have at six or seven. This could produce difficulties at later stages and could stifle spontaneous interest (ibid.).

Older children also appear to profit from nondidactic methods. For example, Zammerelli and Bolton (1977) compared 10 to 12 year olds' mathematical concept formation under three learning conditions. Only the pupils in the play conditions advanced to higher levels of abstraction.

"As the same information was available to both 'play' and 'yoked' groups, it is presumably the added autonomy enjoyed by the 'play' group which enabled them to use the information more effectively. A child given play experience can 'programme' his own activities and this may help him to formulate and change his own hypotheses more effectively than the passive recipient of information. In this experiment, therefore, autonomous activity led more rapidly to greater insight and a higher level of understanding" (ibid., p.160).

The DISTAR program had its roots in Bereiter and Engelmann's (1966) experimental preschool and in Becker's behavioural research. Becker (1977) summarises the four assumptions underlying the model and its essential teaching components:

"First, all children can be taught, regardless of their developmental readiness or background. Teaching failure is not excused. Second, learning the basic skills, including logical procedures, is central to intelligent behavior and should be essential to any compensatory education program. Third, disadvantaged youngsters tend to be behind other students in skills needed to succeed in school as they are now structured. Fourth, in order to 'catch up', the disadvantaged must be taught more in the time available than advantaged children. ...

The model contains seven essential teaching components. First, the teaching of general cases is emphasized so that children can generalize to all members of a set after being introduced to only a few members of that set. Second, since people are the primary instruments of instruction, the number of classroom instructors is increased. By adding more teaching aides, more instruction can be given, especially to non-readers. Third, the daily program is carefully structured; when time is allocated according to teaching priorities, everyone knows what to do and when to do it. Fourth, rapid-paced, teacher-directed, small-group instruction is employed as an efficient way to individualize instruction for the non-reader. Fifth, positive approaches are used to secure and maintain student attention, reinforce correct responses, and identify mistakes. Sixth, teaching staffs are carefully trained and supervised to ensure that appropriate skills have been provided and are maintained. Seventh, student progress is monitored by means of biweekly criterion-referenced tests and reports-of-lessons-taught, both of which help to detect problems while there is time to correct them." (ibid., p.p.521/2)

Behaviour Analysis classrooms, stemming from the S-R and reinforcement theories of Skinner, are similarly based. O'Bryan (1975) criticises such rationale on the grounds that it is based on an incomplete scientific conception of children, has a potential destructiveness on the individuality of the child, a narrowness of instructional objectives and a lack of concern for the child's own potential for structuring the learning experience.

The extent of meaningful adult-child verbal interaction is seen as more important in educating young disadvantaged children than the degree of imposed structure (Woodhead, 1976; Tait and Roberts, 1974; Tough, 1973b). Progress in intellectual growth is linked with the affective domain especially gains in confidence (Tait and Roberts, 1974; Pringle, 1974; Beller, 1973). A pragmatic approach is based on Piagetian structuralist theory yet inserts a degree of structure which recognises some hierarchy of skills and takes its cue from the child's cognitive abilities (Plowden, 1967; Brearley et al., 1969; Parry and Archer, 1974; Manning and Sharpe, 1977).

Brainerd notes that in the United States

"the development and testing of Piagetian preschool curricula remains an active area of educational research. ... the trend away from anything resembling formal instruction ... seems to be accelerating, and there seems to be a growing emphasis on play" (Brainerd, 1983 p.60).

The hierarchical nature of learning was defined by Gagne's conditions of learning which integrated Skinnerian conditioning and Gestalt insightful learning. Eight types of learning are identified, each one being a prequisite for the next type in the hierarchy (Gagne, 1970). In ascending order these are signal learning, S-R connections, chaining, verbal associations, discriminations, concepts, rules and, finally, problem solving.

Vygotsky's (1962) stages in concept formation and attainment are generally in accord with the Piagetian view though Vygotsky places greater weight on the role of language and experience. Three stages are isolated: syncretic or random groupings, thinking in complexes and the 'potential concept stage' in which the child can cope with one attribute at a time but is not yet able to manipulate all the attributes at once. When he can, then he has achieved concept mastery. Pseudo-concepts may arise during the second stage. This happens when a child sees superficial similarities without having grasped the full significance of a concept and is often a consequence of rote learning without an understanding of the underlying attributes.

Nelson's stages of mastery of cultural cognitive systems reflect "Movement from relative openness to a closed state and finally, in some cases, to a new openness characterised by a higher-level understanding of the system as a whole and its potential usefulness" (Nelson and Nelson, 1978 p.275). They suggest presenting a child "first with examples of a general rule that enables him to take in more information about the system and its functioning and then sequence more specific rules in such a way that he is able to build up and coordinate the rules..." They ask "what are the consequences of presenting the child with examples that he is not yet ready for?" (ibid., p.270). Intensive investigation is needed to discover the answer (ibid.). Piagetian theory would suggest a resulting failure in the assimilation and accommodation processes. Personal observation suggests that, at any rate for preschoolers, children adroitly distract from, or ignore, challenges beyond their understanding whilst infant school children may show confusion and eventually loss of confidence.

Kendler and Guenther (1980) investigated the validity of both enlargement and differentiation hypotheses by testing five groups of children aged 3, 4, 5, 7 and 18 (N 160) with classification tasks. From the evidence of the results they argue that "two developmentally linked processes influence classificatory behavior: (a) perceptual similarity based on the mechanism of primary stimulus generalisation and (b) the enlargement of symbolic representations" (ibid., p.346). Their evidence is consistent with the dual-stage developmental model expounded by Kendler and Kendler (1975). This assumes two distinct modes of learning: a single unit (associative) mode in which behaviour is controlled by environmental stimuli and a mediational (cognitive) mode in which behaviour is guided by self-generated representational responses (ibid.).

"In essence the dual-stage model of conceptual development suggests that ontogenetic trends in classificatory behavior should reflect a shift from single-unit to mediational functioning. The mechanism of stimulus generalization would guide classificatory behavior in single-unit functioning, and then a transition to the mediational level should take place. By incorporating the enlargement hypothesis of conceptual development (Saltz and Sigel, 1967) into the mediational stage, one would expect that when implicit symbolic representational responses are initially employed to guide classificatory behaviour, only narrow concepts are being tested. With development the concepts gradually become broader" (Kendler and Kendler, 1975 p.228).

Brown and Desforges (1977) support Schaeffer's model of the hierarchial organisation of subskills which is neither exclusively task specific nor exclusively dependent on cognitive reorganisations but "operates on two dimensions, that of stimulus identity (content) and cognitive processes (structures of skills)" (ibid., p.16). As yet information is lacking on the generalisability of skills. Schaeffer's model "allows only that, given an array of tasks, the presence of hierarchical rules can be determined." (ibid.). Fischer (1980) presents a skill theory of cognitive development. "Skill theory treats cognitive development as the construction of hierarchically ordered collections of specific skills, which are defined formally by means of a set-theory description" (ibid., p.477). Fischer claims this theory takes advantage of the insights of both environmentalists (e.g. Skinner) and organismic-orientated psychologists (e.g. Piaget). It thus embodies the implication that "the organism's control of a skill depends on a particular environmental context" (ibid., p.479).

"Skills develop step by step through a series of 10 hierarchical levels divided into three tiers. The tiers specify skills of vastly different types: sensory-motor skills, representational skills, and abstract skills. The levels specify skills of gradually increasing complexity, with a skill at one level built directly on skills from the preceding level. ... The skills at each level are constructed by a person acting on the environment. She performs several actions induced by a specific environmental circumstance, and the way those actions occur in that circumstance provokes her to combine the actions. The person thus combines and differentiates skills from one level to form skills at the next higher level. ... Development is relatively continuous and gradual, and the person is never at the same level for all skills. The development of skills is induced by the environment, and only the skills induced most consistently will typically be at the highest level that the individual is capable of. Unevenness in development is therefore the rule, not the exception. The level of skills ... is limited, however, by the highest level of which the individual is capable. As the individual develops, this highest level increases, and so she can be induced to extend these skills to the new, higher level." (ibid., p.p.479/50).

In skill theory, cognition includes anything that involves the person's controlling sources of variations including emotions, social skills and language (ibid.). It is assumed that these skills can be described effectively and precisely in terms of sets (sensory motor sets, representational sets and abstract sets). However, Fischer acknowledges the problem of defining the boundaries of a set. He concedes thac "the most useful form of set theory may prove to be the theory of fuzzy sets ... which does not require precise definitions of set boundaries" (ibid., p.483). A further caveat needs to be added: "...resourceful subjects often circumvent the investigator's intentions by doing things in their own way, as in translat- from one psychological content-code to another and using products that were not intended" (Guilford, 1982 p.59).

Generally, skills in the same task domain enable a clear prediction of developmental sequence:

"The theory can be used to predict and explain various developmental phenomena including developmental sequences and synchronies, certain effects of the environment on developing skills, individual differences in development, the nature of developmental unevenness, and structural relationships between developing skills" (Fischer, op. cit., p.504).

Central to the use of the theory is task analysis. This requires a behavioural analysis of performance on each specific task. As Fischer remarks, "task analysis is no trivial matter". Six guidelines are presented. These may be summarised as follows:

1. Identify the set required by the skill.
2. Identify the sources of variation the child must control in the skill.
3. Identify the relations between sets that the child must control.
4. Analyse the particular tasks.
5. Assess the minimal task that would demonstrate the skill in question.

6. To predict a developmental sequence, keep all tasks in the sequence within the same task domain.

"One problem that can arise in interpreting particular tasks is that incorrect task analyses in the developmental literature may interfere with determination of what a person actually must do to perform a task" (ibid., p.506).

Fischer finds that developmental synchrony across task domains is seldom high. He finds moderate synchrony contingent on the correlations produced by age. This may be further confounded by environmental factors such as the degree of practice. Though "unevenness in development must be the rule" (ibid., p.513) "many phenomena that are commonly classified as instances of unevenness are in fact microdevelopmental sequences: The unevenness follows the same pattern in virtually all children in a given social group, and it seems actually to arise from differences in the complexity of skills" (ibid., p.515).

Such instances may explain the horizontal decalages observed by Piaget. Fischer proposes that skill and task analysis "should be able to explain most cases of consistent decalage within a skill domain" (ibid., p.516).

Skill theory may "help to integrate such apparently diverse areas as learning, problem solving, social development, language development, and cognitive development" (ibid., p.523). However, Fischer is exercised by two limitations of skill theory:

"...the need for a more powerful definition of skill domains and the need to deal with the processes by which skills are accessed" (ibid., p.524).

Skill theory does not deal directly with accessing, motivation, memory or attention (ibid.). Gelman's (1982) salutary observations on accessing problems are apposite here.

Within the last decade or so information processing techniques have further refined developmental learning theory. Conceptual Learning and Development (CLD) theory is one such information-processing approach.

CLD theory attempts

"to account for concept attainment from the earliest discrimination of an object, event, quality, or relation as one of permanence different from other things through a high level of attainment, such as by a ... college student who can readily classify many examples and nonexamples and can also indicate the defining attributes of the concept" (Sipple and Klausmeier, 1979 p.33).

CLD is concerned with the child's cognitive structure:

"The cognitive structure of the individual at any time includes everything the individual has learned. ... what the individual is capable of learning at any time is related to the mental operations that have emerged and are functional. The products, or outcomes, of learning may be categorized as perceptual information, verbal information and skills, concepts, principles, structures of knowledge including taxonomies, and problem-solving skills, including strategies of learning and of remembering. These products are learned initially, represented internally, organised, and stored in form of images, symbols, meanings, and relations among images, symbols and meanings" (Klausmeier, 1979a p.6).

Four successively higher levels of the same concept are attained in an invariant sequence and this attainment continues for many years; any three successive levels develop concurrently for one or more years. In ascending order these four levels of concept attainment are the concrete, identity, classificatory and formal levels (ibid.).

The protocols of 292 children were examined year by year from Grade 1 to Grade 12 to identify their performance level on each of four concepts. The same levels of different concepts were found to be mastered at different ages. Thus "the age at which individuals become capable of performing the operations and meeting the other conditions necessary for mastering the classificatory and formal levels is related to the content domain of the concept and the abstractness of the examples of the concept. The operations are performed years earlier on concrete examples that can be experienced through the sense organs rather than on examples represented by symbols only" (Sipple and Klausmeier, op. cit., p.85). Interindividual differences in various operations and strategies for organising, storing and retrieving information contribute to wide age differences in the attainment of concept level (ibid.).

From a 4 year longitudinal study of performance on Piagetian and other tasks by three groups of sixty children starting at age 6, 9 and 12 respectively, and which focused on Piaget's concrete operational stage, CLD theorists discerned a complex picture of cognitive growth (Hooper, Swinton and Sipple, 1979). Quantitative increments in one domain facilitated qualitative leaps in another area. The familiar differences between nonconservers and conservers were evident (ibid.).

"While developmental synchrony was more evident in the initial and final assessment years for the younger age cohorts (ranging from 6 to 11 years of age), the transition from preoperational forms of logical reasoning to a full mastery of the logical groupements specified by Piaget as underlying the concrete operations period appears not to be a punctate, discontinuous episode in the child's life. Rather, the acquisition of concrete operations appears to involve an extended developmental process marked by unevenness and 'reciprocal interweaving' among the supposedly unitary logical concept domains." (ibid., p.p.169/70)

The younger children clearly "approached the logical reasoning tasks in a distinctly different way from that shown by their older counterparts or, indeed, by themselves in the later years of the study" (ibid., p.166). However, results based on analyses of Piagetian measures "did not support the concrete operations stage as descriptive of children's actual cognitive development during ages 7 to 11 (Klausmeier, 1979b p.205). Klausmeier concludes that CLD constructs can account for the course of cognitive development between the ages of five and twelve whereas orthodox Piagetian stage constructs cannot adequately do so (ibid.).

Nevertheless, CLD theory appears to be generally in accord with Piagetian theory. However, Brainerd (1979) draws attention to a fundamental distinction between the two theories. Whereas Piagetian theory asserts that cognitive development controls concept acquisition CLD theory treats cognitive development as an abstraction from the concept learning process. "It is essentially a high-altitude description of what results from many individual concept learning events" (ibid., p.230).

The connections between Piagetian structures and children's actual behaviour were examined by Tomlinson-Keasey, Eisert, Kahle, Hardy-Brown and Keasey (1979). The formation of concrete operational thinking skills were traced over a four year period and the changes in thinking were examined for evidence of any underlying organisation or coherence. Sixty-eight urban middle and lower-middle class children composed the original sample; 30 of these had been lost by the end of testing at age nine. Skills of seriation and number in the kindergarten predicted later behaviour in the conservation of weight and volume suggesting "an empirical coherence to the concrete operation skills being tested" (ibid., p.1161). Analysis showed conservation of mass, weight and volume and class inclusion skills to be causally related with coherence at the individual level. Thus, this 4 year study shows "gradually emerging skills that are related. Further, the form of the relationship is such that certain skills precede others as prerequisites (ibid., p.1162). The authors suggest "a family resemblance framework model in which similarities and relationships are extracted from empirical examinations" (ibid., p.1162).

The models produced in the 1970s by information processing provided "specific hypotheses about the cognitive processes and structures involved in solving problems and understanding language" (Greeno, 1980 p,714). Greeno sees the important task for the 1980s as the development of a theory of learning involving detailed analysis of the acquisition of these knowledge structures. This should be applicable to "the analysis of processes by which children acquire knowledge and skill in school instruction" (ibid., p.714).

Osborne (1982), on the other hand, criticises information processing research on the grounds that cognitive learning is regarded as a purely cognitive matter. Little account is taken of the importance of affective variables in learning and performance. Osborne points out that, as much of human learning is unconscious or reactive and susceptible to motivational influences, much of what we learn may be missed by studying learning analytically rather than holistically."Attempting to understand cognitive processes in terms of the formal relations of language is analogous to saying that we think logically or that our life experience is a manifestation of a conceived formalism, rather than the reverse" (ibid., p.331).

Greeno's (1982) riposte denies that his cognitive psychology is as narrow as Osborne makes out. He claims that it captures the intentionality of cognitive processes in the sense of the content of knowledge and in the goal directedness of experience and behaviour. Somewhat unfairly he counterattacks Osborne's appeal to phenomological psychology by labelling its defense as emotional and ideological. Nevertheless, he concedes that

"We can easily recognise some substantive questions for which the methods and concepts of current cognitive psychology are inadequate. The most obvious deficiences are in the domains of emotion, affect and value..." (ibid., p.334).

Bentler (1980) reviews some recent developments in methods used to analyse qualitative data in the study of cognitive development. These causal models for qualitative data include latent attribute models, prediction models and multinomial response models. Unfortunately, although causal modelling appears to have a useful future, "substantive theory and empirical data make demands on this methodology that it cannot currently meet" (ibid., p.101).

Further discussion of information processing research will be presented after brief descriptions of some of its formal models of cognition.

Spada and Kluwe (1980) present two models of thinking and intellectual development. The first is deterministic and corresponds to the findings of Piaget and Inhelder whilst the second model is a variant of the Rasch latent trait model. Both models were applied in experiments with the same balance scale problems. "It can generally be said that neither the deterministic developmental model, nor the probabilistic linear logistic model can be used to describe completely all of the findings" (ibid., p.p.28/29).

J.R. Anderson (1982) proposes a framework for skill acquisition "that includes two major stages in the development of a cognitive skill: a declarative stage in which facts about the skill domain are interpreted and a procedural stage in which the domain knowledge is directly embodied in procedures for performing the skill" (p.369). This theory is based on Anderson's ACT production system in which facts are encoded propositionally and procedures are encoded as productions. "Knowledge compilation is the process by which the skill transits from the declarative stage to the procedural stage" (ibid., p.369). The declarative system can represent abstract facts which can then be processed by the production system. The latter is able

"to shift attention and control in a symbolic way. These basic symbolic capacities are essential to the success of the learning mechanisms. ...The various learning mechanisms are structured around variable use and reference to goal structures. ...In ACT we see how learning and symbolic processes could be synergetic." (ibid., p.404).

Much use is made in the ACT system of goal structures and Anderson suggests that inspection of the goal structure gives understanding of the role of production (ibid.).

Following Anderson's attempts to build an information processing theory with the ACT system Klahr (1980) focused on the construction of production system models to account for children's behaviour in problem solving. The models for each stage of development are determined by task-specific details of the observed behaviour. They analyse the knowledge that is available and used at different age levels in problem-solving tasks together with the behaviour of short and long-term memory. Assumptions about "STM management interact with conflict resolution principles in determining the system's behaviour. What is not clear is precisely how. The only formalism we have for stating these issues and for exploring their implications are the production systems themselves" (ibid., p.159).

Siegler's decision tree methodology assumes that a person compares objects one dimension at a time and does not integrate dimensions (Wilkening and N. Anderson, 1982). N. Anderson's functional measurement was developed to search for integration rules. Wilkening and Anderson compare these two methodologies for the same balance-scale task. Concern is expressed that "rule assessments based on choice data and binary decision trees may misrepresent subjects' knowledge" (ibid., p.218). Task format that restricts subjects to binary choices may mask their true knowledge and conclusions about the underlying processes may be artifacts of the methodology (ibid.).

In contrast, functional measurement is able to detect both integration and nonintegration rules (ibid.). The binary choice format of many Piagetian tasks raises similar objections. For example, when young children are asked to say which row of beads has the greater number judgments are strongly influenced by the length cue though density could also be used. By applying functional measurement methodology

"Cuneo found that even 3-year olds took account of the density cue. Contrary to Piaget's claims, they did not center on either single clue. Instead, they integrated them according to the adding rule, Judged Numerosity = Length + Density. This integration rule could not be uncovered using the Piagetian methodology or even the binary decision tree methodology because both are predicated on nonintegration rules" (ibid., p.233).

This research highlights a major problem in information processing theories viz. the model of cognitive development produced may be an artifact of the system that is used. The status of IP theories is delineated by Groner, Keller and Menz (1980). Five major points of criticism are advanced of computer program models of underlying psychological theories. IP theories are accepted as theoretically clear and precise but computer implementation "does not automatically guarantee the precise theoretical definition of the terms involved" (ibid., p.164). More seriously, the same problem can arise with entire processes. Thirdly, "the problem of formal representation involves the task domain to be modeled.... Even after very careful and subtle task analysis, there is no a priori or formal way to decide for which other task the present formalization can be generalized, or what exactly constitutes an equivalent task" (ibid., p.165).

Their fourth point concerns empirical validity. Where a model is insufficient ad hoc assumptions and decisions are often made in order to use the model for prediction. Finally, IP theorists are too often satisfied with a modest utilisation of the model whereas "if the program were taken seriously as a model of human performance, it would allow for a vast number of empirical predictions. But so long as they are not followed up, logical sufficiency itself is not a sufficient property for a model of human information-processing." (ibid., p.166).

Klahr's (1980) production system model is used to illustrate each of these five points. Groner et al.'s criticisms are reflected in Pascual-Leone's (1980) critique of information processing simulation. He points out that human differences are neither controlled nor systematically compared; nor are systematic comparisons made of performance across types of tasks. "This type of methodology and data base (I refer to human information-processing researchers in general) has not and could not expose the organismic empirical invariants (the stages and equilibration problems)..." (ibid., p.281). The reasons are clear:

"Fine-grained' step-by-step data repeatedly obtained from a single task with one or more subjects (such that no subject is ever tested across tasks) must necessarily generate structural invariants which reflect constraints belonging to the type of task - for no developmental or human differences organismic constraint is allowed to emerge as an empirical invariant." (ibid., p.281).

Consequently, this methodological strait-jacket results in presuppositions not being properly tested and "taken by many as true by fiat". So "information-processing psychology may, by fiat, manage to repress Piaget's empirical problems of stages and equilibration" (ibid., p.281). Moreover, production systems such as Klahr's (1980 ,) ignore "those aspects in the Piagetian data which justify the admittedly fuzzy Piagetian model of equilibration. ...in particular three salient types of data, ... 'truly novel performances', 'cognitive-conflict situations' and 'motivational or affective paradoxes'" (ibid., p.283).

Pascual-Leone proposes a methodology wherein "organismic factors appear as constructs inferred across types of problems (tasks) and across types of subjects, and which are capable of (constructively) explaining these data in terms of organismic rules" (ibid., p.290).

Klahr (op. cit.) claims that many of the desiderata of a developmental theory listed by Pascual-Leone already exist in production systems. He agrees that task variations are needed "in order to determine the task-free properties of the subject and it is also the case that there are likely to be important dimensions of individual variation" (op. cit., p.160). He foresees future research along these lines.

Spada and Kluwe (1980) also call for more attention to be given to the derivation of problem-solving hypotheses. "What good is a good test of bad hypotheses?" (op. cit., p.29).

"Models based on problem-solving hypotheses are interesting tools for analysing cognitive development. They will become truly valuable tools if it is possible to improve the problem-solving hypotheses and to develop less restrictive models to specify and test them." (Spada and Kluwe, 1980 p.30).

Information-processing models are at present limited to surface manifestations (Breslow, 1981). Breslow argues convincingly against Trabasso's criticisms of Piaget's structural theory. Neither "the creation nor the use of a seriated ordering by subjects in Trabasso's tasks provides evidence against the structural developmental account of transitivity" (ibid., p.348). Breslow points out the hazardous nature of interparadigm criticisms:

"The information-processing approach has been primarily concerned with short-term temporal phenomena in the form of the real-time psychological processes involved in performance on different tasks. Structural developmental theory, in contrast, has been concerned with pervasive, abstract structures that apply to a broad range of tasks (although not necessarily synchronously) and to long-term temporal phenomena in the form of structural changes. Given the different foci of these two paradigms, it is easy for interparadigm debate to go astray by assimilating the concepts of one paradigm to those of another. ...

In view of the predominant concern of information-processing theorists with the formulation of performance models for particular empirical tasks, it is understandable how they might come to identify a concept, such as transitivity, with a given task even when that task is found to be solvable without the use of the concept in question. The main focus of Piaget's theory, however, is on the child's concepts and not on particular tasks per se" (ibid., p.348).

Summary

Piagetian theory asserts that the thought of the child is qualitatively different from that of the adult. Through assimilation and accommodation fresh cognitive structures are created at increasingly higher levels. Crucial to these processes are the child's active involvement with people and objects. Development takes place in an invariant sequence and conflict, the impetus to restructuring, is only experienced when the child has reached a favourable level of development.

Although Piaget's theory of equilibration is criticised for vagueness and lack of empirical evidence no other theory has provided an acceptable alternative. Orthodox Piagetian theory has been revised and expanded by Piaget himself and is constantly subject to refinements by others.

Whereas the child's own propensity for restructuring is central to Piagetian theory, where teaching methods are based on behaviourism the child's autonomy takes second place to rigidly controlled, direct instruction. Head Start and Follow Through reports suggested that the Behavioral Analysis Model and DISTAR programs were the most successful. However, flaws in evaluation procedures lead to the conclusion that chances of later success are raised by preschool programs but no one program model demonstrates superiority. Teacher behaviour is suggested as the key variable together with community support and later school experiences.

Hierarchical models of learning are proposed by Gagné and Vygotsky and by CLD theory. The latter posits that four successively higher levels of the same concept are attained in an invariant sequence.

Skill theory treats cognitive development as the construction of hierarchically ordered collections of specific skills. Its use is contingent on accurate task analysis; little is known about the generalisability of skills.

Information processing has produced precise theories of cognitive processes but the empirical predictions are seldom followed-up. IP models may be artifacts of the system used: until systematic comparisons are made of human differences (as opposed to task differences) production system models will remain incapable of explaining data in terms of long-term structural changes.

In sum, we have only limited knowledge of the complexities of cognitive development. Whatever theory is used due consideration must be given to individual variations in cognitive organisation, motivation and affective factors.

Section B. Language Development in Young Children

It is widely accepted that language has a crucial role in human intellectual development but the precise nature of this role remains unclear. At one end of a polar argument is the view that language results from cognitive development and at the other end the view that language operates on cognition. This may be illustrated by the well-known problem presented by Eskimo children.

"Eskimo children may learn to distinguish seven different kinds of snow through their nonlinguistic experience and theu later learn the labels for them; or, hearing snow called seven different things may prompt them to make distinctions that a child learning English does not" (Hoff-Ginsburg and Shatz, 1982).

An extreme version of the contention that language shapes cognition is the Sapir-Whorf hypothesis which posits that each language imposes a particular 'world-view' on the speaker. Challengers of this hypothesis assert that Sapir's evidence illustrates that language differences reflect the culture and not the reverse (Clark and Clark, 1977). Languages reflect in their vocabularies the needs and interests of their speakers (Stubbs, 1976). This argument will not be pursued here; neither will the arguments concerning universal grammar and biological endowment for, as Chomsky himself says, "the theory of u.g. is not a study of the general properties of language but a postulated component of genetic endowment" (Chomsky, 1981, p.233). Modern syntactic theory (i.e. the transformational grammar of the 50s and 60s) is not per se a processing theory (Marslen-Wilson and Tyler, 1981).

Cognitive theorists adopt variations of the assumption that language presupposes certain underlying cognitive reasons for its use (Clark, 1977; Nelson, K., 1977; Slobin, 1972; Bruner, 1966; Piaget, 1955 for example). Pylyshyn (1977a) speculates that the child possesses "an initial vocabulary of concepts even before it has any recognizable overt linguistic expression" (p.170) and proposes a preverbal system of concepts formed on the basis of potential action as well as potential expression (Pylyshyn, 1979b). Macnamara (1977a) points out that such an assumption logically leads to the position that in order to grasp the meaning of one's native language one needs to know another language already viz. the abstract language of cognition i.e. an innate language in which ideas can be expressed.

Sapir appears to have had similar difficulties:

"It is, indeed in the highest degree likely that language is an instrument originally put to uses lower than the conceptual plane and that thought arises as a refined interpretation of its content. The product grown, in other words, with the instrument, and thought may be no more conceivable, in its genesis and daily practice, without speech than is mathematical reasoning practical without the lever of an appropriate mathematical symbolism." (Bruner, 1966 p.51).

"One word more as to the relation of language and thought. The point of view we have developed does not by any means preclude the possibility of the growth of speech being in a high degree dependent on the development of thought. ... We can see this complex process of the interaction of language and thought actually taking place under our eyes. The instrument makes possible the product, the product refines the instrument." (ibid., p.51)

The underlying notion of reality being represented in cognition before the acquisition of a parallel lexicon also disturbs Olson (1977) and others.

Psycholinguistic evidence on the cognition-language relationship is equivocal. Evidence that verbal labelling (even with nonsense words) assists young children in tasks is provided by Luria and his associates. Labels act as cues in isolating the invariant attributes of different stimulus situations (Luria, 1961; Vygotsky, 1962). Having the name of the concept and the names of its defining attributes facilitates concept acquisition (Klausmeier, 1979b). Blank (1968) notes the lack of concensus on the definition of 'concept' in experimental studies. She doubts that the full range of children's abilities are tapped in work confined to a narrow range of phenomena. No account is taken of children's creative thinking. Viewing language in terms of labels restricts concepts to the classification of objects and thus fails to explain the higher level concepts.

Bruner (1966) reports that if the task is completely in the iconic mode the effect of labelling is cancelled out. It is only when the task is more complex that language is a necessary aid to reconstruction. It is not language per se that reorders experience but rather the tool of language enables the child to check a mismatch between what he sees and what he knows resulting in a genuine reconstructing of how he perceives (ibid.).

Data from two experimental studies suggest that, in addition to visual information "young children can and do make use of other sources of information associated with pictures to augment picture memory performance" (Luszcz and Bacharach, 1980 p.159). Preschoolers will spontaneously use verbal as well as non-verbal information to make decisions in picture recognition tasks but linguistic information may interfere with recognition when two pictures correspond to the same label (ibid.). In these experiments target and distractor pictures either had the same common name or were referred to by different names. Picture memory was found to be attenuated when the targets and distractors shared common names.

"The decrement in performance on nominally identical relative to unique items suggests that excessive reliance on verbal codes at retrieval can result in an inaccurate assessment of memory" (ibid., p.158).

This work of Luszcz and Bacharach seems to be

"an instance where very young children are relying on verbal mediators on a cognitive task that ostensibly requires only visual information. As Anglin has postulated, it is naive to continue to assume that young children in the process of acquiring language do not exploit this rapidly developing competence in executing many cognitive tasks" (ibid., p.159).

It is suggested that adults modify their naming practices when talking to very young children and that the labels they use influence the child's cognitive organisation (Hoff-Ginsburg, 1982). As Brown (1973) suggested, such labels alert the child to similarities and differences which might otherwise have gone unnoticed. Hoff-Ginsburg and Shatz (op. cit.) give this example of linguistic input as an organising influence:

"The first time a baby picks up a quarter and hears the label 'money' it is unlikely that she or he has any notion of what money or quarters are. In providing the label 'money' the mother is anticipating the level of categorization that will first be useful to the child" (ibid., p.19).

Bryant (1974) is critical of 'verbal mediation' theories and cites Kuenne's experiment with size transposition as an example of an inferential conclusion that language is causal when all the experiment proves is that it is associative.

Within the Piagetian system cognitive operations in the young child emerge and develop independently of language.

"Therefore, it would seem to be a paradox to postulate the independence of language and thought, and then to rely on language to infer the existence of certain kinds of thought. If, in fact, thought is not necessarily dependent on language, then it would seem that nonverbal methods should serve as the only appropriate test of Piagetian theory. ... It may in fact be true that certain critical features of thought in the Piagetian system can be assessed only by linguistic methods ... but if this is the case, the need to use language as part of the measurement techniques makes it difficult to establish distinctly cognitive operations, apart from language" (Siegel, 1978 p.47).

From a perusal of Piagetian literature in conjunction with their own data from conservation and language tasks involving the comprehension and production of passives, with 72 children ranging in age from 3:5 to 8:11, Moore and Harris (1978) conclude that there is

"scant evidence for the contention that language skills are a reflection of more general cognitive operations. Constant exposure to linguistic data affords children the opportunity to begin language learning early. They accomplish the task relatively quickly - by means of various inductive procedures and perceptual strategies, the nature of which is the focus of much ongoing research. If the origins of these learning skills exist within more general cognitive operations, it remains for Piagetian theory to stipulate what these universal operations are, and how they are involved in the particular case of language learning" (ibid., p.p.149/50).

That linguistic experience may play a more central role in the acquisition of logical reasoning than Piaget proposed is suggested by Russell's (1982) data from experiments with children in the age range five to ten years. Investigating children's discriminations between (a) necessarily true versus contingent propositions and (b) necessarily false versus contingent propositions Russell finds that "the appreciation of necessary falsity is far easier than that of necessary truth". He suggests that the child's awareness of what cannot be said moulds his conception of what cannot be denied or negated. This involves understanding of what the linguistic system cannot allow (ibid.).

"What we require is a theory which focuses on the linguistic system, not as a mere vehicle for the transmission of logical laws, but as an object of reflection" (ibid., p.265).

Harris (1982) evaluates the evidence for cognitive prerequisition for language acquisition. Firstly, the analogies sought, between children's two and three word combinations and the structures of sensory-motor intelligence, by Brown (1973) and others are considered. The second approach "has attempted to move beyond the level of analogy by establishing a temporal link between the development of sensory-motor intelligence and the acquisition of language" (Harris, 1982 p.189). However, even where "substantial correlations have been found between some aspect of cognitive development..., there are severe problems of interpretation. Such correlations might indicate that cognitive development causes language acquisition, but at least two other interpretations are possible" (ibid., p.189). It may be that some third factor, such as parental stimulation, might accelerate performance in several domains or variations in skills common to both language and cognition (e.g. attention) could produce a correlation between cognitive development and language production (ibid.).

The cognitive prerequisites hypothesis offers no explanation for the different rates of acquisition of language comprehension and production (ibid.). Evidence, especially that from experiments with novel combinations, suggests that

"the precocity of comprehension relative to production cannot be explained by saying that the infant merely understands combinations that he has repeatedly heard, but produces entirely new combinations. Creativity is present both in comprehension and production" (ibid., p.190).

Harris suggests that "whatever cognitive requisites there are have little impact on the timetable for language acquisition" (ibid., p.191). He posits three major factors, more or less independent of any particular level of cognitive development, which probably determine the course of language acquisition. Firstly, linguistic complexity will determine the order in which constructions are acquired. Secondly, it is likely that "the non-verbal context provides an unambiguous gloss on some utterances but not on others" and thus, this "is not a reflection of the language acquirer's cognitive status; it is a reflection of the ease with which the listener can guess what is being said, without understanding all the actual words that are spoken" (ibid., p.192). Thirdly, the asymmetry between comprehension and production "illustrates the fact that what governs the timetable for language acquisition is not the emergence of ideas or 'semantic intentions' per se, but the ease with which such ideas can be mapped into words and vice versa" (ibid., p.192).

This last position is in contrast to Slobin's (1972) contention that "the child intends a range of semantic relations before he commands the full means for expressing these relations" (Slobin, 1972 p.200).

The dilemmas inherent in 'chicken or egg' theories of the relation between language and cognition may be somewhat less of a problem if two further questions are asked. Do language and cognition develop as autonomous systems? Does the nature of the interaction between language and cognition alter with age and growth?

McNeil (1982) denies that language and cognition systems are autonomous. Whereas Bruner's (1977) theory of isomorphism is founded on parallels between actions and language structures, McNeil's Conceptual Basis of Language Theory (CBL) regards language performance as an action and language structure as a wholly separate dimension. CBL bases language acts directly on concepts. McNeil admits that it is not yet possible to say how actions are incorporated but when an act is over the result is a linguistic object that fits into the system of language. If there is a residue of acts not based on concepts then these must be sought from other sources e.g. the pragmatic function. CBL posits, as the basic language act, the presenting of one thought X in the form of another thought Y. Thus XY is the conceptual basis of action, investigating and guiding it to its end point. Both X and Y may derive from culturally constituted categorical knowledge but the combination of X and Y can be unique in the experience of the individual. Grammatical structure is exclusively a property of the linguistic system into which the final product of the act can fit (ibid.). This is in direct opposition to Chomskyan theory.

Unlike Vygotsky, McNeil does not accept two separate lines of development in infancy (a system of representation and a system of speech motor control) which meet at some point. CBL conceives of a fusion of these two lines in a single action control system. This fusion creates a new form of action (ibid.).

Rodgon (1978), on the other hand, stresses that for both theoretical and methodological reasons it is important to first separate, and then consider in relation to each other, two issues involved in the relationship of semantics and overt action. One issue, on the sensory-motor and symbolic levels, is

"the connection between the use of overt actions and the eventual ability to understand and manipulate symbols. The second involves the expression of action relations within the symbolic medium, (which) would indicate the extent to which the child notices action, tries to conceptualize action in symbolic terms, or believes that it is important to comment about action." (ibid., p.71).

From her study of three very young children Rodgon concludes that "action-utterance co-occurence" and the use of action semantic relations are very closely related but are not identical. However there are many differences from child to child and he attests the need for further study of these separate but interrelated issues.

Bruner (1977) suggests that language shapes, augments and even supersedes the child's earlier modes of processing information.

"Iconic representation, in the beginning, is built upon a perceptual organisation that is tied to 'point-at-able' spatioqualitative properties of events."

"Once language becomes a medium for the translation of experience, there is a progressive release from immediacy. ... it permits productive, combinatorial operations in the absence of what is represented." (ibid., p.217)

In Bruner's (1966) view language and experience maintain an important independence of each other until such time as language can be used as an instrument to scale higher levels. Although language can be used for labelling at the outset it is not until after preparation of experience and mental operations that it can be fitted to thought, particularly languages' sentential or syntactic structure. He agrees with Chomsky that "sentences have a compelling power to control both thought and action (ibid., p.55).

In short, Bruner (1966) asserts that when the limits of direct imagery are reached it is necessary to use another means - symbolic representation. A similar view is held by Vygotsky (1962).

Neither language nor thought "can be captured in vacuo as pure form" (Rommetveit, 1978, p.34). From his analysis of eight year old's understanding of simple instructions and situations, Rommetveit adduces evidence of "definite patterns of dependency of thought upon semantic competence which - if further explored - possibly may help us to gain some more insight into the child's thought on premises of its own" (ibid., p.34).

Semantic development and reasoning are examined by Harris (1974) and Carr (1979). Harris describes three experiments with children between 5 and 7 years

"which show that young children can draw reasonable inferences about an entity which has been defined by simple predication. They infer that the entity, so defined, possesses attributes common to members of the superordinate category and lacks attributes which are not possessed. Secondly, subjects have an implicit notion of a class and its members. Thus although a member is assumed to have the attributes of its class, it is not identified with other members of its class unless it shares the appropriate distinguishing attributes" (ibid., p.151).

Children were presented with a nonsense syllable (NS) which was assigned to a category such as man, flower, drink. They were then questioned about its properties e.g. 'a mib is a man. Does a mib eat food? Does a mib have wings? Dop is a red drink. Is dop milk? How do you know dop isn't milk?' "The child's denial that an NS defined as a red drink is milk corresponds to an application of 'modus tollendo tollens' (e.g. if milk, then white; not white (i.e. red); therefore not milk). Clearly these propositions may... be guided by the use of mental imagery. However, Inhelder and Piaget (1958) argue that the symbolic manipulation of propositions does not emerge until adolescence. ...(These) results show that where the material is familiar young children can make valid inferences" (ibid., p.152). If provided with the relevant information young children are able to draw appropriate conclusions though they are unlikely to spontaneously deploy rules of inference (ibid.). (A minor quibble with this study concerns the experimenter's prior assumption that a robin is perceived as a 'red bird'. He was surprised that most children denied "that an NS defined as a 'red bird' was a 'robin'".)

Judgements about the acceptability of anomalous and non-anomalous sentences were elicited from children between the ages of 2 and 5 by Carr (1979). Sentences used in the study were analysed on the basis of whether they had been judged to be anomalous or semantically acceptable. Carr concludes that the development

"of children's capacities to make semantic judgements is irregular and complex, and an explanation is required which can cope with the features of developmental change which individual longitudinal study of children has uncovered. Judgements clearly do not simply reflect linguistic competence. The transition from Stage I through II to stage III seems to represent an attempt by young children to integrate developing language with developing logical notions and expanding experience" (ibid., p.241).

It appears that the youngest children attempted to interpret sentences by using a verification strategy i.e. relating sentence-meaning to their experience and thus checking its factuality. For instance, a "child with a pet cat might agree that a cat could sleep, but assert that a dog could not" (ibid., p.237). During Stage II children apparently regress by accepting almost all sentences, correct and anomalous. This is seen as cognitive change in that "relationships they could not know to be real WERE real... affirming that anomalous semantic relationships which are therefore necessarily unknown are acceptable" (ibid., p.239). By Stage III children "are once again basing their judgements of sentences on their experience, now greatly expanded, and bringing to bear on it fully worked-out ideas of 'known', 'unknown' and 'undecided'. The notion of 'indecision' does not seem to be fully appreciated until Stage III" (ibid., p.240).

Though Carr's thesis seems fairly plausible her methodology is flawed. The experimenter's judgements of errors seem highly debatable. In a sample questionnaire, which was given to a two year old, instances of responses which were deemed incorrect are:

Q. Can fish be tied in a parcel? A. Yes.

Q. Can a rabbit get torn? A. Yes, get their heads off.

Q. Can a ball go and hide? A. Yes.

The only acceptable answer to 'Can a flower wake up?' is considered to be 'No' and it is assumed that dolls should necessarily be regarded as inanimate by very young children.

The study seems altogether over-ambitious. Even adults, as Carr concedes, do not operate simply on the basis of metalinguistic awareness when asked metalinguistic questions. Therefore to attempt to separate young children's metalinguistic from their pragmatic judgements is a Herculean task. How does one allow for the highly imaginative child or the child with an acute sense of humour or for familial idiosyncrasy in the playful use of figures of speech?

Gopnik (1982) notes that many early words encode aspects of plans i.e. actions performed in order to bring about events. Her data suggest that the very young child acquires the words 'gone', 'down', and especially 'there', 'oh dear', 'no' and 'more' at the time he does because they are relevant to his understanding of plans.

From the results of her experiments with preschool children and from many other studies Siegel (1978) concludes that

"there is a strong suggestion of the independence of language and thought for the younger children, and for many of the older ones, in these studies. In the early stages of development of these concepts, there is ample evidence of this independence, but the area of overlap of language and thought increases as the child gets older. Many of the children could not (or did not) use externally given linguistic cues, verbalise the concept, or associate a word with a concept, yet they were well able to demonstrate through nonverbal means the presence of concepts. ... Language skills are not completely correlated with cognitive development, but, more important, this relationship changes with age. Thought cannot be inferred from the language of a young child, although it may be from that of an older child" (op. cit., p.54).

These results seem to confirm the Piagetian dictum that a child can act correctly before he can express that action in words.

Most observers tend to emphasise the dependence of linguistic advance upon cognitive advances for the period between 6 months and 2½ years (Nelson and Bonvillian, 1978). When considering cognition-language interrelationships across the span from about 2½ to 7 years of age the MUTUAL influences of advancing language and cognitive skills are widely acknowledged (ibid.). For example, Hudson, Guthrie and Santilli (1982) demonstrate a relationship between children's cognitive development and their semantic knowledge. They find that 'conservers' offer explanations for their choices in 'more' and 'less' tasks which show a clear understanding of both sequential and simultaneous comparison usage whereas this is comparatively rare with non-conservers. For example, 62.5% of 16 conservers gave a simul~~t~~taneous comparative explanation as compared with 18.2% of 22 non-conservers and 43.8% of 16 transitionals. Hudson et al.'s data differentiate between "kindergarteners' 'more' and 'less' choices derived from linguistic knowledge and those determined primarily by non-linguistic understanding" (ibid., p.137).

Using 60 children between 3 and 7 years of age Trosborg (1981) found a cognitive basis for comprehension of the conjunctions 'before' and 'after'.

"At the level below reversibility, children regard two events described in the sentence as separate entities, neither of which provides a reference point for the other. Reversible thinking was decisive for correct performance on $Before_2$ and $After_2$ constructions. With the achievement of reversible operations as envisaged in the performance of conservation tasks, the children became capable of making inverse-order judgements corresponding to the succession of events" (ibid., p.401).

Siegel (1978) notes a number of studies which find differences between conservers and non-conservers in their understanding of relational terms.

Evaluating the research evidence Keats and Keats (1978) observe that it is not surprising

"in view of the wide range and large number of studies which have been carried out that there is a great deal of conflicting evidence dealing with the language-thought relationship. ...Jenkins (1969) humorously posed a multiple choice question along the lines of, 'What is the relationship between language and thought? : (1) Language is dependent on thought (2) Thought is language (3) Thought is dependent on language (4) None of the above. Or perhaps, all of the above.' ... The correct answer according to Jenkins is, 'All of the above', and it is true that some support can be found for nearly every type of relationship" (ibid., p.166).

An integral part of this relationship is the way children organise their concepts. Katherine Nelson (1977) posits two levels of concept organisation: one personal and cognitive and the other social and linguistic. The child must learn to match the first with the second i.e. his own core concept meaning to the narrow linguistic meaning. She suggests that learning linguistic terms increases conceptual stability and also makes communication about one's concepts possible (ibid.). "Understanding the child's semantic development requires understanding the development of his knowledge system, that is, of the structure of relations within and between concepts." (Nelson, 1978).

Nelson (1978) believes that early word meanings contain chiefly functional information. This contrasts with Eve Clark's (1973) view regarding the importance of perceptual information in concept formation. Support for the latter viewpoint is provided by Tomikawa and Dodd (1980) and by Prawat and Wildfong (1980). Results of experiments with 2 and 3 year olds lead Tomikawa and Dodd to conclude that early conceptualisations and word meanings are perceptually based when perceptual and functional features are independently available. Their first experiment showed that in a sorting task perceptual cues were more frequently used than functional ones. When the particularly salient cues were removed for six other children perceptual similarity still predominated as the criterion for sorting. In their third experiment twelve children were divided into two groups. All the children in the first group met the criterion of learning the nonsense syllable (in a short story) for toys which looked alike whereas none of the second group, where the NS substituted for toys that functioned alike, met the criterion. This result was confirmed by the fourth experiment where 4 out of 6 children met the criterion in perceptual learning conditions but none did in functional learning conditions. Their final experiment to see how children sorted familiar objects for which the perceptual and functional features were independent produced results consistent with the earlier experiments.

Prawat and Wildfong (op. cit.) used four groups of 10 children ages 3, 5, 7 and 8 years respectively and asked them to describe containers shown in different contexts. The results together with the children's comments suggest a developmental sequence which contrasts with Nelson's (op. cit.) theory. The preschoolers seemed to respond according to the form of the object and the intermediate ages according to its use. This led to the conclusion that children's early word meanings are not functional but perceptual. However, this could have happened because the younger children were less consistent in their responding (ibid.). So this possibility was checked by examining their responses to unambiguous prototypical stimuli with the result that it is "safe to say that our data reflect real developmental differences in cognitive processing" (ibid., p.1060).

Eve Clark's studies (1977, 1980; Clark and Clark, 1977) are concerned with the child's strategies for assigning meanings to words in the course of acquisition. She suggests that

"the cognitive principles that the child applies in organising his knowledge appear to provide the basis for his strategies of 'interpretation' when his mapping of difficult words is still far from complete." (Clark and Clark, 1977 p.148)

Children may spend several years adjusting their initial hypotheses about word meanings to the adult meanings. In the early period of language acquisition overextensions are the commonest form of overlaps with adult meanings though underextension may represent the very first stage in the acquisition of a new word. A further possibility is that the child's first meaning does not overlap at all with the adult meaning (Clark and Clark, 1977). Anglin (1978) observes that there are strong biases in the literature "in favour of overextension and against underextension when the only source of information is the child's spontaneous production" (ibid., p.246). In the majority of cases in Anglin's studies when children undergeneralised the problem appeared "to be conceptual in origin rather than as a result of perceptual confusion. ...they could identify a shoe as a <u>shoe</u> but denied that it was CLOTHING..." (ibid., pp.249/50). Whereas overgeneralisation appears to be perceptual in origin: "the noninstance appears to be similar in shape or over-all appearance to an instance of the concept overgeneralized" (ibid., p.250).

Clark's Semantic Features Hypothesis delineates a fixed order of acquisition of components of word meanings, starting with the most general features and moving towards more specific ones. The first dimensional adjectives 'big' and 'small' or 'little' appear in speech between two and three years followed later, in this order, by 'tall-short', 'long-short', 'high-low', 'thick-thin', wide-narrow', 'deep-shallow' (ibid.).

The order of acquisition reflects the relative semantic complexity of dimensional terms fairly closely. (See Bierwisch, 1970). 'Big/small' are used initially to refer to extent on any dimension until the specific dimensional reference is acquired (Richards, 1979a).

This order of acquisition is reflected in a number of studies: for example, Campbell and Wales (1970) observed 'big/wee' used initially with reference to almost all differences in size; Eilers, Oller and Ellington (1974) found 'big/little', 'long/short', 'wide/narrow' acquired in that order; Carey (1978) observed 'tall, wide, fat, high, long, thick,' etc. initially represented as 'big' and their opposites as 'little'.

Clark's top-to-bottom hypothesis has received consistent support (Richards, 1979b). "In a hierarchically organised lexical domain, the order of acquisition appears to be from the top of the hierarchy downward (ibid., p.39). However, Richards points out that confirmation for this hypothesis has thus far been limited to the domain of dimensional adjectives of spatial reference. He cites Anglin's data on the acquisition of concrete nouns which "demonstrate that the top-to-bottom order of acquisition for hierarchically organised systems of semantic features is not a general developmental phenomenon" (ibid., p.39). Nouns at an intermediate level of generality are frequently learned first. "For instance, the child typically learns the term <u>dog</u> before <u>collie</u> or <u>animal</u>, <u>flower</u> before <u>rose</u> or <u>plant</u>, <u>car</u> before <u>Volkswagen</u> or <u>vehicle</u>, and so on" (ibid., p.39).

Assessing the usage of noun level in adult speech Blewitt (1983) also finds that the basic level nouns (e.g. dog) are more frequently used than subordinate (e.g. collie) or superordinate (e.g. animal) levels. Moreover, the frequency of the basic level nouns is increased significantly when adult speech is directed at young children (ibid.). Blewitt's findings support Anglin's position:

"...because of the ways in which parents will name these objects for him he will initially learn terms often at some intermediate level of generality which classify them in maximally useful ways in the sense that they group together objects toward which he should behave in the same way" (Anglin, 1977 p.245).

The first prepositions children use spontaneously to denote relations in space are 'in' and 'on'. Clark (1977) found that when performing linguistic tasks such as 'Put A in/on/under B' children under three show a preference for putting things <u>in</u> a container, otherwise, where B is a flat surface, for putting things <u>on</u> it. In acquiring dimensional adjectives and locative prepositions children generally seem to understand positive terms like 'high', 'long', 'more', 'above', 'before', 'in front of' and 'near' earlier than negative ones like 'low', 'short', 'less', 'below', 'after', 'behind' and 'far' (Clark and Clark, 1977).

These unmarked (positive) terms[1] are thought to be perceptually salient. H. Clark's thesis is that the space in front of a person is optimal for visual perception. Young children pay more attention to height than width; they can deal with vertical asymmetries more easily than horizontal ones and orient most strongly to the upper end (Clark and Clark, 1977).

---

[1] In linguistic theory the criteria for deciding markedness is as follows:

1. Where there is a morphological difference between the marked and unmarked instances, the unmarked instance will invariably be the morphologically simpler of the two.
2. Generally speaking, where the unmarked term <u>dominates</u> the marked term in a semantic field, the unmarked term will be the generic term incorpor-

ating the marked term in certain usages: the unmarked term will therefore be both an antonym and a hyponym of the marked term.

3. Where the semantic opposition is one that refers to the presence or absence of a specific attribute, the unmarked term will always be the one denoting absence. (E.g. same/different, more/less).

4. Marked and unmarked word-pairs will have slightly different syntactic distribution, and will not in all cases be in perfect paradigmatic opposition.

5. Where the marked/unmarked distinction applies to an adjective pair denoting poles of a continuum, the noun denoting the continuum will always be derived from the unmarked term. (E.g. length).

(abridged from Sinha and Walkerdine, 1974 p.3)

The child's a priori preferences for the greater amount, extent and vertical extension make it easier for him to map positive than negative terms (Clark, 1977). Not only are the preferred positive terms the easiest for children to acquire but they are also more often used by adults than the negative terms such as 'backward', 'below', 'out of', 'from'. Hence they are maintained better and with greater accuracy (Clark and Clark, 1977).

Affective polarity is possibly another factor involved in semantic development. This was originally hypothesised by Osgood who identifies 'long' and 'clean' as affectively positive and 'short' and 'dirty' as affectively negative (Richards, 1979b). Osgood and Hoosain report that adults also take less time to comprehend the meaning of affective positive than affective negative terms (ibid.). It would be interesting to speculate on the parallels between the markedness hypothesis and the polarity of young children's ambivalent emotions. To the very young child people, events and objects are seen alternatively as 'good' or 'bad', 'hated' or 'loved'. The swing from one extreme to the other can be very rapid according to the child's feelings at any particular moment. (See Isaacs, 1933; Erikson, 1967; Winnicott, 1974.) There is a strong tendency for orientation towards positive or 'good' objects. Being taller and larger (and therefore stronger and more powerful) figures prominently in young children's aspirations. It seems possible that, in addition to Clark's (op. cit.) explanations of young children's preferences for positive terms, they are also influenced by their emotional states.

Returning to semantic studies, Donaldson and McGarrigle (1974) conceive of

"local rules as being contingent upon predispositions to structure or interpret the world in particular ways which are 'natural' for human beings - or at least which are common in our culture - and which are powerfully present early in life. What appears to happen as children grow older is that the purely linguistic constraints on the interpretation of utterances become stronger and the local rules have a correspondingly reduced part to play. About the process by which this occurs, however, almost nothing is known" (ibid., p.194).

Clark's marking hypothesis has not received consistent support. Evidence that unmarked terms are acquired earlier than marked, or negative, terms is seen for example, in studies by Donaldson and Balfour (1968), Donaldson and Wales (1970), Donaldson and McGarrigle (1974) and Hudson et al. (1982). In each of these studies it is noted, for instance, that 'more' is acquired and understood earlier than 'less'.

Other researchers find evidence of young children's orientation towards and earlier understanding of certain unmarked terms but they do not find that this holds true for all unmarked terms. For example, Levine and Carey (1982) observe that young children comprehend 'before' in the spatial sense earlier than 'after' but, contrary to Clark's evidence, they find that the word 'back' is understood earlier than 'front'. Kuczaj and Maratsos (1975) attest to a simultaneous general acquisition of the use of 'front and back'.

Eilers et al. (1974) find no significant differences with regard to markedness but note individual differences in strategies amongst their rather small groups of two and three year olds. While some children clearly focused on unmarked adjectives, a large proportion showed marked focusing strategies. Eilers et al. suggest that Clark's Semantic Features Hypothesis needs refining in order to take account of those children who demonstrate a strategy of focusing on marked adjectives (ibid.). It would be interesting to know if these same children are affectively oriented towards negativism. Olson (1977) queries whether dimensions have a psychological reality for very young children. He suggests that only absolute qualities are recognised at first and it is only when the child can represent these qualities in terms of quantities that we have the beginning of dimensions.

Even in studies that do report a significant superiority of positive terms

> "there is little evidence to support a consistent acquisition order of learning positive terms before negative terms... So it seems very likely that some children learn the positive term of a polar opposite pair first, whereas others learn the negative term first" (Coker, 1978 p.275).

The results of investigations of temporal and spatial relations with a series of cognitive and linguistic tasks, carried out with 64 children whose ages ranged from 1½ to 3½ and 30 children aged 5 to 7 years when tested in the Bristol Longitudinal Study of Language Development, point to the following conclusions:

"These results call into serious question the theory that the meanings of relational terms (or any lexical item, for that matter) are acquired by addition one by one of semantic features which are identical to those constructed by the linguist as a structural description of the adult lexicon. ... Even if one wishes to speak in terms of semantic components, these will be different in a formal description of the lexicon of any individual child, from those which are derived from a description of the adult lexicon" (Sinha and Walkerdine, 1974 p.30).

Children gradually acquire

"contextual rules which determine socially appropriate usages. ...Each of these strategies represents an approximation to the appropriate contextual rules, within the constraints imposed by his memory-span, perceptual processing skills and level of cognitive development. Since the child's knowledge of the rules will depend upon the contexts within which he has encountered the word, it is reasonable to suppose that for some words, there will be considerable individual differences with respect to the strategies employed, given that children's experiences are not uniform" (ibid., p.30).

Much of the evidence which contradicts Clark's markedness hypothesis is centred on spatial and temporal forms of 'before' and 'after'. Evidence from Durkin (1978) indicates that temporal 'after' is understood better than temporal 'before'. Whereas Feagans (1980) finds temporal 'before' to be acquired before temporal 'after' her data suggest that spatial 'after' is acquired earlier than spatial 'before'. The latter finding is contrary to Clark's temporal-spatial model. Feagans suggests two possible explanations: pragmatically, it is likely that 'after' is used conversationally more often than 'before'. Her alternative explanation is based on the notion that most adults describe temporal 'before' as spatially 'after'. Thus she predicts that 'after' in the spatial sense will be acquired prior to 'before' in the spatial sense by young children (ibid.). This explanation and the notion on which it is based is debatable. However, Trosborg's (1982) data offer some support in that the "ability to seriate objects in space was significantly related to the child's performance on 'before' and 'after' sentences" (ibid., p.397).

Cox (1979) finds that 'behind' is understood earlier than 'in front of'. She points out that, whereas 'front' and 'back' refer to opposite sides of a mid-point, the action, of putting one object in front of or behind another, takes place in front of the child. She surmises that when object A is placed behind object B it is then further towards the positive end of the dimension and therefore is conceptually easier for the child. She suggests that in such cases 'behind' should be regarded as positive.

Kuczaj and Maratsos (1975) find the following order in the acquisition of a mature understanding of 'front', 'back' and 'side':

Level I Child knows the front and back of his own body

Level II Child knows the fronts and backs of various fronted-object types

Level III Child can place a standard in front of and in back of fronted objects

Level IV Child knows the side of his own body and the sides of fronted objects

Level V Child can place a standard in front of, in back of, and on the side of nonfronted objects.

Durkin (1978) also highlights the complexity of 'in front of' and 'behind': they are contingent on the respective situations of speaker and hearer, the reference points and the features of the object.

The ordering of sentences containing 'before' and 'after' conjunctions pose problems for young children. Sentence ordering can also confound research into children's understanding of the temporal terms 'before' and 'after'. Clark acknowledges this complicating factor (Clark and Clark, 1977).

French and Brown (1977) note that 'before' and 'after' sequences are generally easier to comprehend when the described events are logically constrained. When there is no logical order to the events described preschoolers find the interpretation of 'before' and 'after' more difficult. Kavanaugh (1978) confirms this and reports that young children have particular difficulty with sentences where the main clause precedes 'after'. He notes that his findings are not inconsistent with Clark's hypothesis. Weil and Stenning (1978) confirm that a sentence of the type 'A before B' is the easiest (e.g. She brushed her teeth before she went to bed.) followed by 'Before B, A', then 'After A, B' and the hardest type 'B after A'.

Trosborg (1982) observes that 'before' and 'after' sentences present considerable difficulty to young children when they have to rely solely on these conjunctions for their interpretation of event order. Her findings suggest that

> "comprehension of these terms is dependent on the context in which they are presented (arbitrary/logical sentences), upon syntactic and thematic complexity, upon memory load involved by the sentence content, and upon task variables - more so, it seems, than an inherent difficulty of the two conjunctions as proposed by Clark (1971). In fact task variables might even change around the order of difficulty..." (ibid., p.401).

Using sentences which were designed to control syntactic strategy effectiveness Goodz (1982) finds that 'after' is no more difficult to understand than 'before'. Similarly, Coker (1978) finds "no evidence that before is learned before after. On Task 1, the superior performance on after was primarily due to children using the next-event-in-time strategy. ... Similarly on Task 3, the superior performance on before was primarily due to children using the main-clause-first strategy. ... And on Task 2, in which no strategy was possible, there was no difference between before and after" (op. cit., p.275).

The results of the experiments by Goodz

> "strongly suggest that attempts to explain the acquisition of before and after, and presumably other lexical items as well, must take into account not only the semantic features which are thought to make up the meaning of words, but also the cognitive and syntactic contexts in which the words occur" (Goodz, 1982 p.825).

Though widely accepted that children acquire an understanding of 'more' earlier than they do of 'less' there are somewhat different proposals for the immature lexical entry of 'less'. (Donaldson and Balfour, 1968; Donaldson and Wales, 1970; Clark, 1977; Clark and Clark, 1977; Carey, 1978; Laxon, 1981; Hudson et al., 1982). For example, Donaldson and Balfour, proposed that 'less' was understood as synonomous with 'more'. This assumes that children have already acquired the adult meaning of 'more'.

Clark (1977) suggests that young children combine partial semantic knowledge ('more' and 'less' both mean 'amount X') with a nonlinguistic strategy of picking the object with the greater extension or amount. This response happens to coincide with the meaning of the positive term 'more'.

In an investigation of very young children's judgements of quantity, in response to questions or instructions involving the words 'more' and 'same', Laxon (1981) finds "no evidence to suggest that more is easier than same". The findings "suggest that though children may well learn less later than more, it may be because they get by very well by using ± more and ± same to express ± polarity. More and less may well be used as synonyms because the antonym for + more may well be - more or even + same... but a failure to distinguish between more and less does not mean that - polarity is not operating under another label..." (ibid., p.540).

Carey (1978) suggests that "there is no point during development when less is represented in the lexicon as a word with a meaning that is not in opposition to more" (p.129). She also suggests that this holds true for other polar adjectives. Partial meanings for comparative spatial adjectives persist because, although the binary distinction of polarity seems to be represented from the beginning, the distinctions among dimensions such as length and depth are much more complex. Thus when learning new adjectives the polarity features are already available as lexical organisers. Usually, in comprehension, the dimension in question is obvious from the context and in production the child can rely on the more general 'big' and 'little' if not sure which word applies to a particular dimension.

Hudson et al. (1982) differentiate between children's choices which are derived from linguistic knowledge and those derived from nonlinguistic understanding. Their data show the majority of wrong 'less' choices appearing to be based on nonlinguistic information.

Clark (1980) is also exercised by the role of non-linguistic strategies. Using children aged 1:6 to 5:0 she examined the role of nonlinguistic strategies in the acquisition of certain orientational terms. Clark concludes that:

> "Very young children show a strong nonlinguistic preference for selecting the upper or topmost surfaces of objects, a strategy that produces consistently correct responses to top and, equally consistently, incorrect responses to bottom, front, and back. In order to distinguish the operation of such a nonlinguistic strategy from actual knowledge of the term top, it is necessary to wait until the children contrast top with bottom. Similarly, knowledge of front and back can only be attributed to children once they contrast these terms with top and bottom as well as with each other. Children's nonlinguistic preferences and strategies, then, are in one sense what determines both order of acquisition and relative complexity of meaning. Where strategy and meaning coincide, the meaning should be relatively simple to acquire. Where they do not, children have to relinquish their nonlinguistic strategies before they can work out the adult meanings, and this is what makes those meanings relatively complex." (ibid., p.337/8).

Further evidence of children's non-linguistic strategies is reported by Wilcox and Palermo (1982), Hoogenraad, Grieve, Baldwin and Campbell (1978), Donaldson (1978), Durkin (1978), Harris, Macrae and Bassett (1978), Macnamara (1977a), Townsend and Erb (1975) and Donaldson and McGarrigle (1974).

It is now evident that word meanings are only one source of information among many others that young children use to interpret utterances. Younger children rely more on nonlinguistic information than do older children. Between two and six years the youngest children rely relatively little on word meanings per se (Wilcox and Palermo, 1982). Where a child's interpretation of the context of an utterance conflicts with his imperfect semantic knowledge he is more confident of the nonlinguistic context (Hoogenraad et al., 1978). (Nonlinguistic strategies are also discussed in the next chapter.)

It is noted that the immature understanding and use of common prepositions, comparative and quantifying adjectives is by no means confined to children under six. Acquisition is a protracted affair and continues at least through the infant school years (Durkin, 1978, 1981).

The studies already discussed shed some light on children's semantic development and on some of their pragmatic strategies. Karmiloff-Smith's (1978) study is rare, if not unique, in that it focuses on the interplay between simultaneously developing syntactic, semantic, phonological and pragmatic strategies employed by young children.

Early Genevan psycholinguistic research seemed to confirm that lexical and syntactic developments were strictly dependent on cognitive ones (Karmiloff-Smith, 1978). "Whilst Piaget has stressed explicitly the effects of cognition on language development, he has left implicit the effects of the latter on cognitive growth" (ibid., p.18). Current linguistic studies from Geneva are examining some of these issues (Vuyk, 1981b).

Based on experimental results, Karmiloff-Smith posits a model of the interplay of basic and standby procedures during language acquisition. She asks how such a model fits within Piaget's epistemology:

"Although conceptual progress indeed plays a role in language acquisition levels, we should like to place stress on the procedural aspect. Much stress has been placed on Piaget's concept of conservation. We should like to suggest a shift of emphasis from 'conservation attainment' to the more dynamic concept of 'conservation seeking'. Only by seeking to conserve patterns and strategies can the child get a grip on the environment and gain what Nelson (1974) has called increasing 'predictive control'. If the child were to take into consideration each new piece of information and constantly remodel his procedures, he would not have the opportunity of consolidating the procedures in the first place. Clearly conserving linguistic patterns and meanings of expressions plays a decisive role in getting a grip on linguistic inputs. Overgeneralisation is a powerful device for simultaneously simplifying and unifying; this again helps the child to seek patterns which link counter-examples. The fact that with development there appears to be more interplay between procedures may be due to the child's attempt to seek patterns and cohesion within his own multifacet approaches to the linguistic input. Piaget's concept of 'augmentative equilibration' ... may find empirical validation in this direction." (ibid., p.16)

Thus, Karmiloff-Smith proposes that it is neither mere progressive accommodation to the linguistic input nor, solely, communicative intent which pushes the young child into adding morphological markers to his language. She hypothesises that, when the child first introduces, say, articles fairly consistently, "he is doing so for himself and not initially for communicative purposes" (ibid., p.17). In short, the push for the child is the need to get a grip on the linguistic input.

The emphasis in Karmiloff-Smith's paper is on language as a problem area for children in its own right. The proposals stated are working hypotheses for further study (ibid.). Perhaps her approach, which focuses on the simultaneous interplay between syntax, semantics, phonology and pragmatics, points the way for future research.

It is probable that, in a literate society, children's awareness of language structure per se, or treating language as an object, is a consequence of the processes of learning to read and write (Olson and Torrance, 1983). As Olson and Torrance argue, "it has to do with learning to differentiate form from content, what is said from what is meant" (ibid., p.145): but there is no simple causality: what is decisive is the attitude to language, not simply the skills of reading and writing (ibid.). They suggest that there is a shift from "attention to the beliefs and intentions of persons towards the meanings and structures of sentences; a shift from intended meanings to sentence meanings" (ibid., p.148). This attention to language is probably related to "the types of questions teachers ask in school: 'Why do you say that?' 'How do you know?' 'How can you tell?' and so on" (ibid., p.154). In addition to classroom talk and learning to read and write, knowledge of the metalanguage may derive from preschool literacy events and conversations with literate adults (ibid.). (See also the findings of Wells, 1981 described below).

Recent studies of metacognitive development show that explicit information can be an effective way of advancing children's explicit knowledge about communication (Robinson, 1983). Four and five year olds who were told when and why listeners understood or failed to understand what the speaker meant "improved more than those in the practice groups in both level of performance in communication tasks, and in degree of explicit knowledge about message ambiguity and its role in causing communication failure" (ibid., p.130). (The practice groups played the same games as the 'guidance' groups but were provided only with appropriate behaviour modelled by the experimenter when it was her turn to play.) In none of the groups were there any signs of advances in performance only or of advances in explicit knowledge of strategies only. The exception was questioning where there was improvement without there necessarily being associated improvements in explicit knowledge (ibid.).

So far this review of children's language development has been mainly concerned with psycholinguistic studies and theories. Sociolinguistic studies highlight further problems in children's linguistic development and competencies. Cultural deficit theories imply that children from certain social groups, when evaluated according to the schools's definition of what constitutes language proficiency, are found to be deficient. Therefore in order to enable them to fit in with the existing educational process they require remedial instruction. (See Becker, 1977 , for example.)

The degree to which a child achieves mastery of the language system and turns his language competency into a particular standard of performance is, according to Bernstein (1971), a consequence of social structure. He claims that different patterns of social relationships give rise to the use of distinct linguistic codes. These codes generate different social, emotional and intellectual patterns. "A tiny percentage of the population has been given access to the principles of intellectual change whereas the rest have been denied this access" (ibid., p.175).

"The concept 'compensatory education' implies that something is lacking in the family and so in the child. As a result the children are unable to benefit from schools. It follows then that the school has to compensate for the something which is missing in the family and the children become little deficit systems. If only the parents were interested in the goodies we offer; if only they were like middle-class parents, then we could do our job. Once the problem is seen even implicitly in this way, then it becomes appropriate to coin the terms 'cultural deprivation', 'linguistic deprivation' etc.. And then these labels do their own sad work" (Bernstein, 1971, p.192).

Bernstein (1970) argued that the real problem of educability for such children is the confrontation which occurs between (a) the school's universalistic orders of meaning and the social relations which they generate and (b) their own particularistic orders of meaning and the social relations which generate them. Nevertheless, he suggests, there is nothing to prevent a child from internalising and using universalistic meanings as long as his curiosity and capacity for exploration can be utilised.

From a study of English and American teachers' language attitudes Shafer (1978) observes:

"Both English and American teachers displayed clearly formed language attitudes which stereotyped the speech of many of the children they teach as culturally deprived. More that 90% of the English teachers in this study and almost 50% of the American teachers linked the language of children from working-class groups and racial and ethnic minorities as representing a language deficit. ... The English teachers in particular had been reinforced by the early writing of Basil Bernstein in thinking of working-class language as a 'restricted code' and therefore representing cultural deficiency. Both English and American teachers saw in the language of working-class children and in the children of minorities the self-fulfilling prophecy that children who come to school with deficiencies are bound to leave school with deficiencies, because of the difficulties they encounter in school. ... It seems clear that the language attitudes held by many teachers in England and the U.S. result in unconscious stereotyping not only of the speech of working-class children and ethnic and racial minorities in both England and the U.S., but also in the stereotyping of the learning potentials of children who exhibit language not judged by the teachers to be Standard English. Few teachers either in England or the U.S. seem to be aware that there is a logic to non-standard English and that logic has any implication for educational practice." (ibid., p.442).

'Cultural difference' theorists recognise language variety but deny that skill in speaking or thinking is produced exclusively by the customary use of one kind more than another. Language performance in schools and in tests is not necessarily indicative of language competence (Labov, 1969; Tough, 1973a; Rosen, 1972; Stubbs, 1976).

"It is more plausible to account for the effect of language on educational success in terms of people's attitude to language rather than to propose that different language varieties have different cognitive effects." (Stubbs, 1976 p.111)

Different groups of people have different norms of appropriate language use. There is no linguistic justification for judgments concerning the superiority of one variety over another. Nevertheless there is a complex sociolinguistic relationship between a child's language and his success at school. A number of factors operate in this relationship; the most important of which are the form of teaching within our culture and patterns of social stratification (Stubbs, 1976; Fuchs, 1973).

Given the wide range of linguistic diversity Trudgill (1974) argues that there are three options. The first is the traditional solution of trying to eliminate all varieties except 'standard' English. Trudgill shows why this has not worked arguing that it is wrong psychologically, morally and linguistically. Secondly we can attempt a policy of 'bidialectalism'. Torrey (1973) sees this as the way to accommodate cultural variety but Trudgill avows this would only work for written language. Thirdly we can work to change people's attitudes to language so that each has respect for the other's variety.

"We should, according to this approach, teach children the ability to read standard English, but, beyond that, we should simply attempt to educate our society to an understanding, appreciation and tolerance of non-standard dialects as complex, valid and adequate linguistic systems" (Trudgill, 1974 p.82/83).

Meanwhile non-standard speakers continue to be at a disadvantage in the educational system. Therefore the most satisfactory solution, from the point of view of a linguist, is the adoption of a combination of bidialectalism and appreciation of dialect differences providing the former is handled sensitively enough to foster linguistic security (ibid.).

Shields and Steiner (1973) advocate a greater willingness on the part of teachers to accept the child's own language as an initial teaching language until he can switch with ease to standard English. Tough (1973a; 1973b) argues that the young disadvantaged child has not necessarily had different encounters with the environment from the advantaged child but meaning has not been put into them by an adult. Whilst he may differ relatively little from children from more advantaged backgrounds in his language competence, his language performance generally is poorer because he sees no need to use more elaborate forms of speech. Therefore we must find ways of provoking him into using language for different purposes than hitherto (ibid.).

Tough (1970) analysed the speech of three year olds from "advantaging" and "disadvantaging" social environments i.e. children of professional parents and children from homes where both parents had completed their education at the minimal age and where the occupations followed required no particular linguistic skill i.e. amongst unskilled and semi-skilled workers.

In addition, the quality of linguistic fostering in the home was assessed by interview and only those children of above average intelligence were included. Analyses of the speech data recorded in play situations revealed significant differences in the use of language by these two different groups of three year olds. For example, "the analyses of particular components of the language ... seem to indicate that the language of the child from the favouring home is leading towards explicitness, specification and analysis. The child from the less favouring home is seen to use language just as much but he operates over a more limited and limiting range: his language tends to be more implicit, is lacking in specification and generally remains bound by the immediate situation" (ibid., p.25).

Tough (1973b) summarises what three and five year old linguistically disadvantaged children fail to use their language for:

" 1. They rarely report on past experiences.
2. They do not extend the imaginative situation.
3. They do not offer explanations and justifications.
4. They do not predict and plan.
5. They do not consider alternative possibilities.
6. They do not recognise and work for the solution of problems."
(ibid., p.15).

She discusses the educational implications:

"How can we provide an environment which will foster the skill in using language over the wide range of purposes already familiar to the child who comes from the educating home?

There may be some who would suggest that such a policy implies the imposition of middle class values on the child and who would question our right to do this. But what we are talking about are strategies which are induced by using language, strategies which facilitate thinking and learning, strategies which the young child has a propensity to acquire. Surely we have no right to deny any child access to these.

If such strategies are to be learned then the task is to redirect the focus of the child's attention: to place him in situations where what he seeks to do with his language demands using it to represent past happenings, to anticipate, predict and plan, to explain, to analyse, and to search for clues to meaning and for the interpretation of his experiences.

There seems to be no other way of doing this except by the method that some mothers use intuitively, of starting where meaning lies for the child: of helping him to realise that value is set on his thinking, by providing him with an appreciative but provocative audience, an interested adult. ... the adult who is leader and provocateur, an active participant in building up new attitudes in the child, who encourages new interests in what is happening around him, who will set up in the nursery an environment which first engages the interest of all and then invites each child to explore, to question, to solve problems, to learn to use language as a means of learning, who recognises and uses every opportunity to help the child extend his thinking." (ibid., p.p.22/23. See also Tough, 1973a and the various publications of the Schools Council Project 'Communication Skills in Early Childhood'.)

This is very different from the form of intervention advocated by DISTAR enthusiasts as reviewed in Section A. Martlew (1980) observes that where the adults adopt a didactic interventive role, the characteristics of the preschool playgroup children's speech to them "does not seem to be particularly distinguished. Far from encouraging more explanations and suggestions, for instance, they elicited more naming and one word utterances" (ibid., p.203/4). However, the study in question covered only six months of initial preschool experience (ibid.). Martlew's major finding is that, in their initial preschool experience, children could benefit "almost as much from peer interactions as from having a highly interventive programme" (ibid., p.205). The latter consisted of carefully organised, small group activities under adult guidance in which staff maximised their interactions with the children. The matched playgroup children were given the same physical resources but the staff adopted caretaker roles with little intervention (ibid.). Neither of these two regimes appears to have utilised the sensitive and imaginative part of adult-child interaction advocated by Tough (op. cit.).

Robinson (1983) notes that naturalistic studies have begun to identify parental ways of interacting with preschool children which are related to reflective thinking. "Discussions between mother and child about motives, feelings and intentions are considered to be important" (ibid., p.133). However, whether such parental behaviours are causal in promoting reflective thinking in other tasks awaits experimental confirmation (ibid.).

Further investigations of mother-child interactions confirm many of Tough's observations. For example, Mills and Funnell (1983) find a relationship between maternal 'world links' and child deductive reasoning but it is not found for "the ordinary links which all mothers of 2-year-olds make" (ibid., p.165).

"Some mothers mentally situate the child within his own family and social context and take pains to make his current experiences relevant to him. Others do not. All mothers tend to respond when the child, looking at a book or picture, says 'That's a cow' by saying, 'Yes it says moo' or 'Yes there are two of them, do you see the other one? ... But only some mothers would be likely to say, 'Yes, do you remember that weekend when Daddy took us on a picnic and the cow ate your hat?'. Or 'It's the same colour as the cow who fell in the canal - you used to love that story when you were little.'" (ibid., p.165).

Such parents rehearse, activate and elaborate the episodic memories of their children (ibid.).

Some support for Tough's claims concerning language usage is contained in a report drawing on a longitudinal study of language development in 32 children from one to seven years of age (Wells, 1981). This involves data from recordings of spontaneous conversations in the home, interviews with parents, tests and teacher assessments. However, an attempt to apply Tough's categories of language usage to the speech, recorded at 3½ years, of a comparable group of 32 children, "failed to show a clear pattern of class-related differences in the functional use of language. Nevertheless, it is possible that such class associated differences in language use only emerge in the later pre-school years" (ibid., p.192).

To date, Wells finds only a weak relationship between family background and oral ability when this is measured quanticitively.

"However, the stronger claims concern qualitative differences, either of code or dialect. Such class-associated differences, it is argued, lead to an incompatibility between the language habitually used at home and that used at school, which is substantially greater for lower-class children" (ibid., p.192).

The associations between non-standardness of dialect and accent, as assessed at age 5, and attainment at age 7 were relatively small but when assessed by teachers at age 7 a substantial negative association with attainment was found (ibid.). It seems that

"the non-standard dialect or accent that is associated with low attainment is not a source of difficulty that results from a child having learned in the pre-school years a variety of language which is incompatible with that of the school, but rather a mark of group identity which is adopted or exaggerated after entering school, perhaps in response to his growing perception of himself as a low attainer" (ibid., p.p.192/3).

Wells' study reveals no serious incompatibility between "the language spoken in the two settings of home and school,... At most there is some tendency, barely reaching a level of statistical significance, for lower-class children to be less advanced in their mastery of the common language, with socially based variations not being, in themselves, a major problem" (ibid., p.194).

Nevertheless, scores on measures of oral language at successive age-points showed progressively stronger predictions of later educational attainment. Wells claims that

"although an adequate mastery of spoken language is an important prerequisite for progress in school, it is knowledge and ability with respect to written language that is of particular importance.

Two reasons can be suggested for this. First, a major part of the curriculum in the early years of schooling is concerned with the acquisition of literacy, and ability in this area constitutes a large part of any assessment of attainment. Secondly, the skills involved in learning to read and write are characteristic of much of the learning that takes place at school in their relative abstractness and emphasis on the symbolic nature of linguistic representations. ... ...it is one of the chief aims of the school to help the child to 'disembed' his thinking from the supportive context of immediate experience and to bring it under the control of meanings that are encoded in the linguistic message alone. Clearly, learning to use the written language is one very important way of developing this ability" (ibid., p.194).

This complements the arguments of Olson and Torrance (1983).

In the area of literacy Wells finds substantial class-associated differences in the children's preschool experience. It appears that

"one important source of the class-associated differences that emerge in school-based assessments of attainment is the difference between homes in the value that is placed on literacy and in the steps that parents take to transmit this value to their children" (ibid., p.195).

He hypothesises that in homes orientated towards skills of literacy

"there will be a greater likelihood that, when appropriate, the exchange of meanings in conversation will be influenced by the more analytic manipulation of experience that is particularly characteristic of written language (Olson, 1977), with the result that the children have the opportunity to develop an awareness of the way in which language allows particular situations, problems and predicaments to be represented in symbolic categories and relations, which can be communicated about and acted upon independently of their context of origin" (ibid., p.195).

The recordings in the above study did not include the pre-bedtime period, when reading and related talk is most likely to occur.

"Clearly, a study designed specifically to test the hypothesised relationships will be necessary before any firm conclusions can be drawn about the influence of parental literacy on the quality of parent-child conversation in the home and, as a result, on the child's preparedness for formal education" (ibid., p.196).

Wells acknowledges the obvious association between parental literacy and occupational status. He also believes that there are "reasons for doubting the appropriateness of treating the population as if it could be divided into two (or three) homogenous social classes" (ibid., p.193). The Registrar-General's categorisation I - V was used together with the terminal level of the education of both parents.

It is likely that research into the links between children's linguistic abilities and their home experiences will, if based on such crude social classifications, remain problematical. Over a decade ago Rosen (1972) pointed out that it is

"no simple matter of dividing the working class into 'skilled' and 'unskilled', nor of the niceties of the Hall-Jones scale. We must distinguish between those who are quiescent and defeated and those who are articulate, highly verbal, between those who are submerged in what Freire (1971) calls 'the culture of silence' and others who are capable of being quite explicit about principles and operations in those areas of experience which have been their universities" (Rosen, 1972 p.8).

Rosen asks what distinguishes

"the language of Liverpool dockers from that of Durham miners or Clydeside shipbuilders or London railway men or Coventry car-workers? ... There are other strands which need teasing out, like the tradition of certain kinds of nonconformity and the persistance of an oral tradition amongst Irish workers. It takes a Ewan MacColl or Charles Parker to have an ear for such things" (ibid., p.9).

The apparent articulateness of the Welsh in comparison with their English counterparts could be added to Rosen's examples of possible language differences <u>within</u> social classes which await investigation.

Summary

The answer to the age-old question of whether language is dependent on thought or vice versa remains theoretical. There is a great deal of conflicting evidence and some support can be found for almost every type of relationship. Even where substantial correlations have been found between language skills and cognitive development at least three interpretations are possible. A Piagetian view is that "conservation seeking" impels the child to get a grip on linguistic input.

At first over- and, sometimes under-extension of adult word-meanings occur but eventually children adjust their own concept meaning to society's narrow linguistic meaning. Though nouns at an intermediate level of generality are frequently learned first, research evidence consistently supports the view that the order of acquisition of relational terms reflects their relative semantic complexity: for example, 'tall', 'wide', 'thick' etc. are initially represented as 'big' and their opposites as 'little'.

Evidence for Clark's marking hypothesis is more equivocal. Whilst many studies find that positive (unmarked) terms are acquired earlier than negative (marked) terms there is evidence that this does not hold true for all unmarked terms nor for all children. Comprehension is dependent also on context, syntactic complexity, memory load and task variables. Young children combine partial semantic knowledge with non-linguistic strategies.

Mastery of the language system is seen by many as a consequence of social structure. Compensatory education is deemed necessary for those children considered 'linguistically deprived'. An opposing view is taken by 'cultural difference' theorists who deny that skill in thinking or speaking is produced by the customary use of standard English. They see attitude change as necessary in order that non-standard speakers may cease to be handicapped by the educational system. Informed and sensitive adults can foster the widening of language strategies in young children; recent studies are identifying relationships between parent-child interactions and reflective thinking. It may be that parental literacy is a key factor in the quality of parent-child conversations and consequently the child's preparedness for formal education.

CHAPTER 4

Review of Literature (ii): Problems in Testing Concept Mastery

There are many problems in testing procedures. The effects of obvious difficulties such as poor rapport between tester and testee, anxiety, fatigue, coaching, the child's inability to follow directions and the effects of lax administration and scoring procedures have been well documented by such standard texts as Anastasi (1982) and Cronbach (1970).

Test norms which are used for selective or predictive purposes should, ideally, not be biased in favour of any sub-group in the population. Standardised tests have been seen as having a middle-class bias not only in content (Labov, 1969) but also in the advantage the middle-class child is thought to have in the comparative ease with which he can respond to a test situation (Tizard, 1974). However, if the test's purpose is to predict success or failure in our educational system, such a bend in the yardstick may not render it invalid for its particular purpose (Davie et al., 1972). (Though the knock-on effects may be damaging to the children concerned.)

Jensen (1980) protests that 'culture bias' became a popular cliche. Its use was characterised by a "lack of conceptual clarity and objective criteria" and only "recently has the topic of test bias become of major theoretical and empirical interest to psychologists who specialise in psychometrics" (ibid., p.370). Jensen attacks three inadequate concepts of test bias. The first, the "egalitarian fallacy", is

> "based on the gratuitous assumption that all human populations are essentially identical or equal in whatever trait or ability the test purports to measure" (ibid., p.370).

Therefore, according to this definition, the "perfectly nonbiased test ... would reveal reliable individual differences but not reliable (i.e. statistically significant) group differences..." (ibid., p.370).

Secondly, the "culture-bound fallacy" is "based on the content validity (or face validity) of test items. A subjective judgement is made as to the degree to which particular items are 'culture bound'" (ibid., p.371). Such subjective criteria are seldom clearly specified but usually are applied to items that

"involve scholastic or 'bookish' vocabulary or knowledge, or knowledge of fine arts or items that reflect what are imagined to be the moral, ethical, or aesthetic values of the white middle-class" (ibid., p.371).

The fallacy, Jenser avers,

"is not in the possibility that some test items may discriminate between different cultural groups because of the groups' differences in experience but that such items can be identified or graded as to their degree of culture-boundedness merely by casual inspection or subjective judgment" (ibid., p.371).

The third fallacy, attacked by Jensen, is the belief that "because a test was standardized on a given population, it is ipso facto biased or unfair when used in any other population" (ibid., p.372). As Jensen points out, reliability and validity for any group other than the standardisation population cannot be taken for granted but must be investigated in their own right. Simply including minority groups in a standardisation population does not assure lack of bias against any of these groups (ibid.).

"Proper standardization for different subgroups should consist of comparable item selection procedures performed separately within each subgroup. ... Only in the final norming ... for the composite sample should the subgroups be combined in proportion to their numbers in the general population" (ibid., p.373).

Jensen distinguishes between the concepts of bias and unfairness. He uses 'bias' to refer to

"systematic errors in the predictive validity or the construct validity of test scores of individuals that are associated with the individual's group membership" (ibid., p.375).

Thus, psychometric bias is "a set of statistical attributes conjointly of a given test and two or more specified subpopulations" (ibid., p.375). He considers the notion of unfairness not to be an attribute of tests themselves but of the uses to which they are put.

Content, suggests Jensen, may be ordered on a continuum of cultural loading with 'culture free' and 'culture bound' as hypothetical end points.

The relative positions of test items on this continuum can only be determined by subjective judgement. Such judgements, "even when they represent a high degree of concordance among many judges, do not prove the objective correctness of an item's relative position on the culture-loading continuum" (ibid., p.375). Jensen casts doubt on whether any objective set of criteria, for determining cultural loading, could be found.

Psychometric bias, though, can be defined and identified by external or predictive validity criteria and by internal or construct validity criteria (ibid.).

"In terms of predictive validity, a test is defined as biased with respect to two (or more) groups when either the regression error of estimates, or both, are different (i.e. a statistically significant difference) for the two groups. This means that for an unbiased test any given estimated true score based on the test yields the same prediction of criterion performance, with the same degree of accuracy, regardless of the group membership (e.g. race, class, sex) of the person obtaining that score" (ibid., p.454).

Even a biased test, as thus defined, can still achieve fair selection if the predictor variables include the use of a person's group membership and the test has adequate predictive validity <u>within</u> each sub-population (ibid.).

Internal and construct validity criteria are more complex. It is possible, by various statistical methods, to test hypotheses of probable differences if "the test were in fact culturally biased with respect to any two or more groups of interest" (ibid., p.455). Jensen describes methods for determining whether group differences are

"merely a result of a difference in overall level of ability, as compared with true cultural difference. The two cannot be distinguished, of course, if one proposes the purely ad hoc and quite implausible hypothesis that the cultural differences <u>between</u> two population groups perfectly simulate all the 'internal' psychometric effects associated with ability differences <u>within</u> each population group" (ibid., p.455).

(See also Anastasi, 1982; Cole, 1981; Linn and Harnisch, 1981 for discussion of Item Response theory models; Scheuneman, 1979 for a chi square procedure.)

Cole (1981) reviews the approaches taken to detect bias and, like Jensen and others, argues that bias questions are fundamentally validity questions. She believes that no one bias approach should stand alone:

"Various types of information about bias - whether from predictive validity perspectives, methods traditionally associated with construct validation, or judgmental procedures - provide different types of evidence. Combining these types of information gives a more complete understanding of possible bias than using any one type alone" (ibid., p.1068).

Jensen observes that the various item biases tend to cancel one another out in the total score. Where this happens the test should still be regarded as biased from a psychometric standpoint (op. cit.).

Jensen's final conclusion is that

"Most current standardized tests of mental ability yield unbiased measures for all native-born English-speaking segments of American society today, regardless of their sex or their racial and social-class background. The observed mean differences in test scores between various groups are generally not an artifact of the tests themselves, but are attributable to factors that are causally independent of the tests" (ibid., p.740).

Such a conclusion seems over-optimistic. The procedures advocated by Jensen for eradicating test bias are stringent. Whereas some tests may fulfill his criteria it is doubtful, without supporting evidence, that "most current standardised tests" do so. So despite Jensen's defence of standardised tests it is important to be vigilant for signs of the middle-class bias noted earlier and below. Even if empirical proof of bias is missing the possibility of hidden bias needs to be borne in mind.

However psychometrically sound a test may be, young working-class children may have had less practice in the skills required of a test respondent. It has been observed that young working-class children spend a great deal of their time playing outside with peers and siblings and less time than their middle-class peers in situations indoors which involve paying close attention to an adult (Hubbard, 1974; Halsey, 1972; Newson and Newson, 1968). Despite some reported discrepancies in this general finding (Chazan, Laing and Jackson, 1971; Hubbard, 1974) it would appear that the socially disadvantaged child, at any rate, is also at a disadvantage in this respect. Bernstein argues that success at school depends, amongst other things, on "a willingness to undertake solitary privatised activity and to give selective attention to an adult" (Tizard, 1974).

Thus, because of the nature of his preschool experience, not only is a more privileged child more likely to have acquired the skills that the test measures (Jencks, 1973) but he is also, compared with his less privileged peers, able to fall more easily into the role of test respondent (Tizard, 1974).

Experimental evidence which partials out social class variables is reported on the effects of high noise levels in the home on children between 4½ and 6½ who, in the laboratory, were asked to carry out (a) a search task which indicated the distribution of the child's attention between focal and incidental stimuli and (b) a matching task under an auditory distraction condition and in relative quiet (Wohlwill and Heft, 1977). After the inter-correlations between the task variables and the child's age, preschool experience, family's income and other environmental variables had been partialled out, noise level in the home (as assessed from family activity, audio and other electrical appliances together with noise from exterior sources) correlated .20 with the subject's search time ($p < .05$) and -.29 ($p < .01$) with recognition of incidental stimuli. In the second experiment, though the mean performance of the children from the noisiest homes was poorer in the quiet condition than that of children from quieter homes, they were less strongly affected by the distractor. Their mean increase in response time was 2.04 seconds as compared with 6.92 seconds for the low-noise group. Thus

"it appears that their ability to selectively attend to the relevant stimulus features in each situation was adversley affected by the high noise levels in their homes in spite of their apparent adaptation to these conditions" (ibid., p.132).

Experience of testing can lead to increased test sophistication (Anastasi, 1981) and the establishment of response set whereby the context-ual cues the child first encounters influence his responses to subsequent related tasks (Hargreaves, Molloy and Pratt, 1982). As noted in Chapter 2 increased self-confidence and motivation allow children to function at their best level during a test (Zigler and Butterfield, 1968). Some children enter school with abundant self-confidence whilst others may only acquire it gradually; others may never do so or their experience of school may weaken whatever confidence they have. Thus, irrespective of test bias, interpretation of test results requires considerable circumspection.

Allied to these arguments is one advanced by 'cultural difference' theorists (op. cit.) that tests and school-based learning share certain common features viz. they are symbolic, abstract and complex. If tests are demonstrated to have high validity any bias must derive from the criteria (Rosenbach and Mowder, 1981). If the criteria of success are centred in middle-class culture laden schooling where verbal skills predominate then predictive validity studies will not reflect bias in the test (ibid.). The bias is in the criteria not in the test per se. Jensen also points out that "poor criterion measurement can make a good test look bad" (1980, p.383).

Testing for concept attainment poses special problems in addition to those already mentioned. If a test item appears satisfactory can we be certain that a child fails it because he does not understand the concept? Conversely, can we be sure he understands the concept if he passes an item? Does the child perceive the task in the same way as the adult does? Do linguistic differences between the adult and the child influence either the child's response or the adult's interpretation of that response? These questions are inter-related and are fundamental to the testing of concept attainment.

They surface in the controversy about training experiments, acceleration of learning and Piaget's own experimental procedures (See Vuyk, 1981b). Smedslund (1977) protests that Piaget focussed on the child's forms of logic and merely assumed the child's understanding of terms and instructions. Smedslund's position is "to presuppose logicality in the other person and always to treat his understanding of given situations as a matter for empirical study." (p.3).

"A number of investigations subscribe to the view that nonconservation judgments are simply due to language difficulties and explain the acquisition of conservation by the fact that the child grasps conservation once he has mastered the language sufficiently to be able to understand the instructions and explain his reasoning." (Inhelder et al., 1974, p.17/18)

Inhelder is not convinced:

"How does it happen that a particular child uses and understands terms such as 'more' and 'less', 'just as much' and 'the same amount' in the adult sense when he appears to interpret them differently? There can certainly be no question of the difficulty being purely verbal." (ibid., p.100)

CLD theory claims that attainment of successively higher levels of the same concept and the meaning of the word proceeds simultaneously (Klausmeier, op. cit.). A child's meaning of a word is equivalent to his mental construct at each level (ibid.).

Siegel notes a number of studies which show that "the form of the verbal material in questions designed to elicit logical responses can be shown to influence whether a child will be able to give a correct response." (Siegel, 1978 p.46)

Donaldson (1978) points out that the child's interpretation of the words of the speaker is influenced by the expectations which he brings to the situation. When the language of a task conflicts with the child's expectancies in the situation it is the language which gives way (Olson and Torrance, 1983). Conversely, the way a situation is described will affect the way the child construes it as will the adult's non-linguistic behaviour. The influence of the eliciting context of a task is shown by Campbell and Wales (1970), Donaldson and McGarrigle (1974), Hoogenraad et al. (1978), Light, Buckingham and Robbins (1979), Hargreaves et al., (1982), Meadows, (1983), Mills and Funnell (1983) and others. (See Chapter 3)

"...it is necessary to distinguish ... between the child's knowledge of the meaning of a preposition and his ability to make use of this knowledge in different contexts." (Hoogenraad et al., 1978, p.181)

Donaldson (1982) suggests "a mode of thinking and behaving in which language is effectively embedded in the totality of events among which it occurs, a mode in which meaning does not depend upon the words alone" (ibid., p.202). Outside certain formal modes it is possible that "the normal thing is to take the language together with its context - physical and personal - and to try to make sense of the whole". Furthermore, suggests Donaldson, "the younger the child the more likely it becomes that non-linguistic cues will outweigh linguistic ones" (ibid., p.230). It is possible that some children tend to respond to the impersonal, physical features of a task whilst

"others - or the same children at other times - may tend to respond to interpersonal or social ones. Some may concern themselves with level of liquid and length of row, while others wonder about what the experimenter is up to. It is likely that many wonder about both and are apt to be swayed by each in turn" (ibid., p.206).

The 'correct' interpretation of a word on one occasion is no guarantee of full understanding on another (Donaldson, 1978). The child may find it difficult to pay scrupulous attention to the language in its own right or may even be unable to do so. He may amend the question to one more to his own liking. Donaldson cites Inhelder's observation of differences in children's 'attitude' towards the words of the experimenter:

"Children from more privileged backgrounds are more likely to pay scrupulous attention to the words of the question, reflecting on them, analysing them before answering. By contrast, the less privileged children have a strong tendency to substitute a 'more natural' question for the one the experimenter has asked" (ibid., p.89).

It is apparent that when a child is addressed, even if he does not understand the words, he generally feels obliged to respond. Young children not only provide answers to questions that are bizarre: they do the same for questions whose terms' intended meaning may be obscure and for questions which are anomalous in various ways (Hughes and Grieve, 1980; Carey, 1978; Light, 1983; Hoogenraad et al., 1978; Robinson, 1983).

"The gap between questions as adults present them (intended questions) and questions as children respond to them (received questions) is wider than is often supposed" (Hughes and Grieve, 1980, p.149).

This "has significant implications for what we think is happening when we attempt to gauge the young child's cognitive/linguistic abilities by means of the question and answer process" (ibid., p.160).

Even such reputable tests as the WPPSI, the McCarthy (General Cognitive Scale) and the Stanford-Binet (years II - VII) assume that basic concept words in the test directions are understood by the child (Kaufman, 1978). For example, the WPPSI Animal House test alone uses eight basic concepts in the instructions spoken by the examiner. Kaufman cites the following direction with the assumed concepts italicized (underlined): "Fill in the different houses without skipping. Do them one right after the other." (ibid., p.209). The concepts 'without', 'right' and 'other' might also have been highlighted by Kaufman. Nonverbal tasks in the WPSSI and McCarthy Scales are supposed to measure perceptual organisation and visual-motor coordination but "the scores of a child who does poorly because he fails to understand some verbal concepts in the directions do not validly reflect his nonverbal intelligence" (ibid., p.210).

As already noted in the previous chapter, it is apparent that many young children are more confident of nonlinguistic clues than of the utterances addressed to them (Hoogenraad et al., 1978; Clark, 1980). What information is available to the child that makes this possible? Macnamara (1977b) emphasises the importance of gestures and looks to the child in his second year. Data from experiments, with children between two and six years, reported by Wilcox and Palermo (1982) indicate that "a significant number of young children know that: (1) adults are likely to ask them to place objects in their normal relationships, (2) the repetition of a request indicates that the child's prior interpretation was incorrect, (3) different words typically have different meanings, and (4) only certain words go in certain linguistic contexts" (ibid., p.149).

Varying the contexts of items within a test may affect young children's performance. Harris, Macrae and Bassett (1978) found that the older children (6 and 7 years) in their study were highly sensitive to the listener's needs and the younger children (4 and 5 years) did pay attention to the needs of the listener. However, their data indicate that although younger children tailor their communications to fit an initial context they fail to adjust to a new context.

When testing children's understanding of concept terms we need to know if their correct responses reflect understanding or whether their conceptual preferences make it appear that they know these terms before they actually do (Clark, 1980).

We need to be aware of the logical and grammatical constraints of the sentences used to formulate task instructions especially when these incorporate the words 'before' and 'after' (Clark and Clark, 1977; Donaldson, 1978; Durkin, 1978; French and Brown, 1977; Kavanaugh, 1979; Trosborg, 1982; Weil and Stenning, 1978).

The memory load of complex sentences must also be considered especially with young children (Trosborg, 1982; Kavanaugh, 1979; Bryant, 1975). Bryant suggests that a child may fail a task because he cannot remember the premises. However, Bryant's conclusions, drawn from his transitivity experiments, are open to criticism (Vuyk, 1981b).

It is hypothesised that

> "...children understand 'before' well enough to form some internal memory representation of the temporal relation between the events that are described by the before sentences, and therefore when they delete the action described in the subordinate clause it is probably because they take the subordinate as already given. On the other hand, they do not understand after very well, and therefore are more often thrown back on verbatim recall of the sentence they heard, in which case they are more likely to recall only the clause heard more recently" (Weil and Stebbing, 1978, p.406/7).

Young children can ignore syntactic rules. "All the cars in the garages" may be treated as equivalent to "all the garages have cars in them". The question of which noun is quantified may be disregarded (Donaldson and McGarrigle, 1974).

The properties of the materials used in conceptual tasks may also be important variables in children's interpretations. For example, the position in front of an object with a recognisable front (e.g. car, horse) is easier to locate than a similar position when the object has no easily recognisable front (e.g. a cube) (Weil, op. cit.). A young child's interpretation of a task may be influenced by the manner in which he would represent the physical objects if no adult were there (Donaldson, 1978). Young children rely on more external frameworks for their perception and memory of individual objects than adults do (Bryant, 1975).

When using three-dimensional materials masking can facilitate performance on certain tasks by enabling the child to 'decentre' more easily (Flavell, 1974) or by reducing the likelihood of "a perceptual-iconic representation becoming dominant and inhibiting the operation of symbolic processes" (Bruner, 1977, p.213). Two-dimensional materials can remove the masking variable and produce somewhat different results as Cox and Willetts (1982) found when investigating primary schoolchildren's understanding of before-behind and left-right relationships.

Ault, Cromer and Mitchell (1977) tested 39 preschool children enrolled in a Head Start program to see whether their performance on a three-dimensional version of the BTBC would appreciably exceed their scores on a comparable two-dimensional version.

"A comparison of 2-D and 3-D versions of the BTBC indicated that children either seemed to know a concept on both versions or did not identify it correctly on either version above random guessing levels. ... The present data suggest that the BTBC may assess a level of cognitive development at which stimulus dimensionality alone is not an important factor" (ibid., 9.187).

The authors speculate that a "year's experience of the Head Start program could have prepared the children to deal with 2-D stimulsus". They also recognise that "perhaps these children were not deprived of 2-D stimulus from books or television prior to their enrollment (sic.) in Head Start" (ibid., p.187).

Notwithstanding discrepancies reported within and between communities in the UK, evidence suggests that the vast majority of preschool children here are familiar with pictures even within the most disadvantaged populations. (See Hubbard, 1974; Halsey, 1972; Chazan, 1971; Newson, 1968). It is unfortunate that Ault et al. did not test the children early in their Head Start program.

Pictures may also mislead. Poor or ambiguous drawings are not the only source of difficulty. A young child's propensity to 'centre' on a detail which is irrelevant to the task can cause misinterpretation (Vernon, 1968). Pictures which represent another's point of view can confuse the child who is not yet able to 'decentre' (Piaget and Inhelder, 1956).

Whether a test uses objects or pictures the number of articles in an array can affect performance on a task involving size relationships. Young and McPherson (1976) report superior performances of young children when the problems are within their subitizing ranges.

Thus it is apparent that there are a number of factors which can vary the situational context of any test. An even more fundamental problem is posed by the BTBC which is designed to identify "individual children whose overall level of concept mastery is low" (Boehm, 1971). Is a level of concept mastery a viable construct?

Piagetian theory asserts that the answer lies in the level of thinking developed by the child, as ascertained by Piagetian tasks, and, for convenience, delineated in stages and substages. Within the broad stages of preoperational and operational thinking, however, any one child does not necessarily achieve exactly the same level of development across all concepts at any given time (Inhelder et al., 1974; Piaget, 1977a, 1978; Vuyk, 1981a, 1981b).

"... little is known about the 'oblique' relationships between concepts within the concrete operational period, particularly between the first notions of conservation of discontinuous quantities and those of continuous quantities (liquids, solid matter and length), which are generally acquired between the ages of six and nine years" (Inhelder, 1974 p.247).

"The problem of the developmental links between the various concepts is evident in the 'decalages', i.e. non-synchronous acquisitions of concepts which are based on the same operatory structures" (ibid., p.15).

Jensen cites evidence from many correlational studies and from his own principal components analysis of the intercorrelations between fourteen

Piagetian tasks to support his claim that

"... the general factor of the Piagetian battery is almost pure g in the Spearman sense. Vernon (1965) factor analysed a large number of Piagetian tests along with many standard IQ tests and found a large general factor - unquestionably Spearman's g - common to both types of tests." (Jensen, 1980 p.675).

Jensen notes that Vernon's study is almost "the only wholly correct factor analysis of Piagetian tests that (he has) found in the literature" (ibid., p.675). He points out that researchers who orthogonally rotate factors, by use of computer programmes such as Kaiser's varimax, prohibit the emergence of the large general factor. He insists that, in such studies, "either oblique rotation should be done to permit the hierarchical extraction of g, or the g factor should be extracted (as the first principal factor) prior to rotation of the remaining factors" (ibid., p.675).

Operativity (or intelligence as defined by Piaget) is identified as a distinct factor in five year old children by Lunzer, Wilkinson and Dolan (1976). A range of Piagetian and language measures together with two standard tests of intelligence and tests of short and long-term memory were given to 210 children aged from 5:6 to 6:0. Seven factors were distinguished with operativity emerging as Factor I and language as Factor II. The same seven factors appeared in both the oblique and the varimax solutions. However,

"although language and operativity remain factorially distinct even at the second-order level, these factors are oblique and quite strongly correlated with each other" (ibid., p.292).

In view of these factor analyses, and others, it seems unlikely that a test of conceptual development will be unidimensional. Is it, then, possible to use a multidimensional test to measure an overall level of concept mastery?

CLD theory suggests that any of the four successive levels of concept attainment, identified by Klausmeier and his associates (1979) could be ascertained for any particular concept but the same levels of different concepts are mastered at different ages. This is related to the content domain of the concepts and the abstractness of the examples that are presented to the child. Therefore, according to CLD theory, it would appear necessary to predetermine the level of content domain and the abstractness of each item by sophisticated empirical means, such as information processing techniques, before one could construct a test which yielded an accurate assessment of a level of concept mastery.

The necessary operational structures do not emerge at the same time for all kinds of concepts (Klausmeier, op. cit.). For example, concrete, three dimensional examples are understood earlier than symbolically expressed concepts. According to CLD theory, individual concepts move through various levels, primarily as a result of learning, at differing rates and they co-exist at different levels of sophistication (Brainerd, 1979).

Vygotsky's (1962) hierarchical model of concept organisation suggests that concept maturity is not reached until the child can manipulate all attributes at once. A stage level might be ascertained. For example, Vygotsky's 'potential concept stage' could be presumed when a child can respond correctly to items which require the manipulation of only one attribute at a time but cannot cope with items which require manipulation of several attributes.

Clearly, in none of the above models is concept mastery regarded as a unidimensional construct.

Can a summary score, obtained from a multidimensional test with a content that is made up of several domains which are apparently independent of each other, estimate with any accuracy a child's level of concept mastery?

Recent developments in the analysis of item response patterns suggest that global scores can be misleading indicators of a child's attributes (Harnisch, 1983; Hutt, 1980; Hunt and Lansman, 1975). Their limitations can be seen by considering their mathematical foundation.

"The same number-right score can be obtained in many different ways. Even on a short test of just five items, a score of three can be obtained by ten different combinations of correct and incorrect responses. The number of possible combinations expands rapidly with increases in test length" (Harnisch, 1983 p.191).

"More often than not, the scores resulting from the performance on complex tasks required in psychological tests have been assumed to represent comparable composites in different individuals, and error variances have usually been disregarded" (Hutt, 1980 p.241).

"In the past, many psychologists have seen individual differences as sources of error variance to be eliminated through the use of efficient experimental designs. ... There is a real danger that the laws revealed by these mean scores are not true of any individual subject, but are the result of averaging two or more distinct, nonmodal patterns of behavior" (Hunt and Lansman, 1975 p.107).

Though summary scores may serve as useful predictors of success the interrelatedness of subsets of items make the diagnosis of individual strengths and weaknesses difficult. Ways of improving on subjective judgements in diagnosis are suggested. Analysis of item response patterns based on Student-Problem (S-P) theory is advocated by Harnisch (op. cit.) whilst microdiagnosis based on an experimental-clinical model is proposed by Hutt (op. cit.). Hunt and Lansman suggest that experimental psychologists should, after determining group trends, "go back and see to what extent these trends are typical of individual subjects, and that where they are not, we seek explanations" (op. cit., p.107).

As part of a screening programme Spector (1979) administered the BTBC to more than 300 kindergarten children living in a middle-class community. "By examining the types of errors and questioning children about their responses after the test sessions, several possible reasons for failure to respond to concept items can be added to the reasons Boehm offers" (Spector, 1979, p.564/5).

"Cognitive factors that appear to contribute to a lack of comprehension of basic concepts include (1) inability to focus on the key words in the directions, (2) complexity of directions, (3) deficits in spatial perception, (4) lack of knowledge of concept labels of vocabulary deficits, (5) difficult level of abstraction, (6) difficulty with negative concepts and (7) inadequate auditory memory for sentences.

A general deficit in cognitive functioning appears to affect the child's ability to comprehend basic concepts." (ibid., p.567).

Thus the test results of children with scores below average on the BTBC should be analysed for "possible error patterns that may indicate specific problems. ... A child who does poorly on this test of concepts should be suspect, and further diagnostic testing should be initiated to determine the extent of the child's deficits" (ibid., p.567).

Whilst acknowledging their limitations it is possible that global scores obtained from a multidimensional achievement test may indicate the standard of performance reached by groups of pupils, relative to other groups, where the criteria are defined in terms of particular knowledge or skills. This is contingent on the criterion measurements being reliable and unbiased. (See Jensen, 1980) A child who obtains an optimal score on a criterion-referenced test which yields an 'all-or-none' score for each item is assumed to have reached some preestablished level of mastery (Anastasi, 1982).

"In the construction of such tests, two important questions are: (1) How many items must be used for reliable assessment of each of the specific instructional objectives covered by the test? (2) What proportion of items must be correct for the reliable establishment of mastery? In much current testing, these two questions have been answered by judgmental decisions. Substantial progress has been made, however, in developing appropriate statistical techniques that may provide objective, empirical answers" (Anastasi, 1982 p.96).

If the criterion is concept mastery then a clear definition of 'mastery' is a necessary step in constructing criterion measurements. Theory suggests that the optimum scorer should, for example,

(a) be able to manipulate all the attributes of the concepts at once (Vygotsky, 1962),

(b) understand the concept at the formal level, as posited by CLD theory, whereby its meaning is the same as society's meaning and the defining attributes can be named (Klausmeier, 1979 a,b),

(c) be able to use the concept in generalisations and reversible operations (Piaget, 1952, 1955).

Though the concept of reversibility is ubiquitous, in its various forms it is found at all stages of development (Moore and Harris, 1978). Great care would be needed in drawing conclusions from a child's naming of the defining properties of a concept. Anglin observed that

"Although in discussing the words children were sometimes able to mention the properties which adults use to define them, there were indications that they may not have realised that those properties were defining. ... Such observations are consistent with the hypothesis that even though children often know a considerable amount about a given concept, including the properties which adults use to define it, they may not realize that these properties are defining or in other words the criteria which determine what objects are instances and what objects are noninstances of that concept" (op. cit., p.p.253/4).

Total scores which are less than one hundred per cent correct may indicate the degree to which the subject has achieved 'mastery' however defined. A three-way distinction may be employed such as mastery, non-mastery and an intermediate or doubtful interval (Anastasi, 1982). On the other hand, as the literature on the use of global scores indicates, the averaging of a number of discrete item scores, may mislead. Where the mastery of concepts is concerned, theory suggests that each concept or, at the least, each group of concepts being tested, should be examined separately. Each concept or test item may need its own criterion reference.

Anglin postulates

"A continuum of concepts from primitive to scientific with the former being vague, concrete, undifferentiated, and completely lacking coordination between extension and intension and with the latter being precise, abstract, differentiated, and characterized by a complete coordination between extension and intension. On such a continuum the child's concepts would occupy a region closer to the primitive pole than to the scientific pole and the adult's concepts would be somewhere in the middle" (1977 , p.264).

CHAPTER 5

Review of Literature (iii):

The Boehm Test of Basic Concepts

A The Theoretical Framework of the BTBC.

B Reviews of the BTBC.

C Bias and the BTBC.

D Validity of the BTBC.

E The BTBC and Children with Sensory or Learning Disabilities.

A The Theoretical Framework of the Boehm Test of Basic Concepts

The theoretical foundation on which the Boehm Test of Basic Concepts (BTBC) was based is not made explicit in the test manual (1971) nor does the literature reveal any elaboration of theory by Boehm. Statements about the purpose of the test appear rooted in pragmatism:

"The BTBC is designed to measure children's mastery of concepts considered necessary for achievement in the first years of school." (Boehm, 1971 p.3)

"Every child who enters a school system comes with a different background of experience, and thus starts out with a different body of knowledge and set of understandings. It is essential that educators focus their attention on such variations in cognitive development, with respect to their nature, origins, modifiability, and implications for future achievement." (ibid., p.3)

Elsewhere, Boehm defines "basic concepts" as a group of concepts "which involve the ability to make relational judgments either between objects, persons or situations, or to a standard. ...they all involve judgments which apply to shifting situations" (Boehm, 1980 p.3).

Boehm (1971) writes of "concept mastery" and suggests remediation instruction based on sequential steps in concept presentation. She directs that

"remedial instruction should start with a presentation of the concepts in an elementary context before more subtle or complex connotations of the concepts are presented" (ibid., p.14).

This direction, together with the suggestions for classroom presentation of concepts, appears to be based on sequencing from single concrete examples to extensions. This is in contrast to the view that early cognitive learning moves from relatively general ideas to increasing specificity (Nelson and Nelson, 1978).

No mention is made of the Piagetian view that concept acquisition progresses from the child's 'action' (however broadly defined) to conceptualisation. Piagetian theories about preoperational thinking, transitivity, cognitive conflict and structural change are ignored by Boehm (1980, 1971, 1966). The BTBC is not a test of concrete operational thinking per se.

A The Theoretical Framework of the Boehm Test of Basic Concepts

The theoretical foundation on which the Boehm Test of Basic Concepts (BTBC) was based is not made explicit in the test manual (1971) nor does the literature reveal any elaboration of theory by Boehm. Statements about the purpose of the test appear rooted in pragmatism:

"The BTBC is designed to measure children's mastery of concepts considered necessary for achievement in the first years of school." (Boehm, 1971 p.3)

"Every child who enters a school system comes with a different background of experience, and thus starts out with a different body of knowledge and set of understandings. It is essential that educators focus their attention on such variations in cognitive development, with respect to their nature, origins, modifiability, and implications for future achievement." (ibid., p.3)

Elsewhere, Boehm defines "basic concepts" as a group of concepts "which involve the ability to make relational judgments either between objects, persons or situations, or to a standard. ...they all involve judgments which apply to shifting situations" (Boehm, 1980 p.3).

Boehm (1971) writes of "concept mastery" and suggests remediation instruction based on sequential steps in concept presentation. She directs that

"remedial instruction should start with a presentation of the concepts in an elementary context before more subtle or complex connotations of the concepts are presented" (ibid., p.14).

This direction, together with the suggestions for classroom presentation of concepts, appears to be based on sequencing from single concrete examples to extensions. This is in contrast to the view that early cognitive learning moves from relatively general ideas to increasing specificity (Nelson and Nelson, 1978).

No mention is made of the Piagetian view that concept acquisition progresses from the child's 'action' (however broadly defined) to conceptualisation. Piagetian theories about preoperational thinking, transitivity, cognitive conflict and structural change are ignored by Boehm (1980, 1971, 1966). The BTBC is not a test of concrete operational thinking per se.

Nevertheless, certain items involve some understanding of either equivalence, ordination, seriation, number invariance, quantity relations or classification.

Boehm (1971) emphasises 'presentation' by the teacher. This gives the impression that direct instructional methods based on behaviourism are recommended. Indeed, if "the class average on the test is low ... various group training procedures should be considered" (ibid., p.12).

"The assumption that children have mastered the basic concepts necessary for understanding and following directions by the time of school entry needs to be questioned. Indeed, studies of reasonably typical samples have shown that as many as 60 out of 100 children entering kindergarten may be unable to mark the right end of a line, or to indicate the area below a pictured table" (ibid., p.3).

Boehm considers such findings "justifiably a source of concern because of their possible implications for initial and hence subsequent school achievement" (ibid., p.3). She clearly regards these findings as indicative of early cultural deprivation. Whether the test or curricula on which it is based might be culturally irrelevant to these children is not even considered.

Thus the BTBC is based on a 'cultural deficit' theory of development. "The instrument is designed to assess beginning school children's knowledge of frequently used basic concepts... (ibid., p.4) and "any concept understood by less that 100 per cent of the children in a class should, ideally, be considered a problem for remediation" (ibid., p.14). So, when evaluated according to the educational system's notion of which concepts and their labels should be understood, those children failing such a criterion are to be regarded as 'deficient'.

It is unclear from the BTBC manual whether or not a distinction is made between concept understanding and concept labelling. The test's rationale and purpose is described as though the two were synonymous or, at any rate, synchronous. Its purpose is to identify "children whose overall level of concept mastery is low" and also "individual concepts with which large numbers of children in a class may be unfamiliar" (ibid., p.4).

However, when discussing reasons for concept difficulty, Boehm makes reference to the distinctions:

"Other difficulties which may contribute to incorrect responses on the BTBC are:
(a) inherent difficulty with the concept
(b) language difficulty where the specific label is not known or alternative concept labels are unfamiliar, or
(c) pictorial representation that is ambiguous or out of the child's realm of experience." (ibid., p.14)

No mention is made of the Piagetian view that a child can act correctly before he is able to put into words what he has just done; nor of the pseudo-concepts that can arise from learning a concept label without fully understanding the concept. (See Vygotsky, 1962)

In an earlier study Boehm (1966) acknowledges that concept difficulty may be due to a true absence of concepts. 'Same-different' emerged as one of the most difficult concepts at kindergarten level.

"Does an incorrect response on such a concept indicate that the concept was not understood or that the test pictures were ambiguous, or that a language difficulty was present that could have been avoided had the items been worded with alternative labels such as 'alike and not alike'?" (Boehm, 1966 p.89).

Additionally, she notes that

"Piaget also pointed out that when mathematical concepts are imposed prematurely, the resultant learning may be of the name rather than the essential idea" (ibid., p.89).

She does not, however, raise this last point in the test manual.

Underpinning the rationale of the BTBC appears to be the assumption that there is a convergence of the child's linguistic and cognitive abilities which is tapped by the test. The linguistic abilities employed must include semantic, syntactic, phonological and pragmatic aspects but these considerations are not examined by Boehm (1971, 1980).

In sum, the BTBC appears to have been constructed on a pragmatic basis. Underpinning this seem to be the following theoretical assumptions:

1. Linguistic and cognitive abilities converge.
2. At a given point in development this convergence can be tapped and measured.

3. The resulting measurement assesses concept mastery.
4. Early mastery of certain concepts is necessary for school achievement.
5. Failure to respond correctly to a proportion of items testing these concepts indicates cultural deprivation and subsequent educational failure.
6. The schools' curricula are correct and relevant to all children.
7. Efforts at remediation should be based on behavioural techniques.
8. Scant attention need be paid to structuralist theories of cognitive development.
9. Children's concept mastery is related to their socio-economic status.

Each of these assumptions will be discussed in detail later.

It is possible that the apparent vagueness of Boehm's theoretical stance on conceptual development stems from her pragmatism: certain children were seen to have difficulties in understanding directions; therefore they needed identifying so that their 'faults' could be remedied.

B <u>Reviews of the BTBC</u>

The BTBC has been reviewed by Dahl (1973), Freeman (1972), Lawlor (1972), McCandless (1972), Noll (1972), Proger (1972) and Smock (1972).

The purpose of the BTBC "as stated by the author is 'to assess beginning school children's knowledge of frequently used basic concepts widely but sometimes mistakenly assumed to be familiar to children at their time of entry into kindergarten or first grade.'" (McCandless, 1972 p.335) McCandless considers this statement "accurate, modest, and realistic" though he recognises that the "initial item content was empirically and apparently somewhat subjectively determined" (ibid., p.335). "No validity evidence other than face validity is presented, although the face validity is convincing enough" (ibid., p.335).

Smock endorses the content validity as adequate: "...the concepts selected were important for the understanding and following of directions and occurred frequently in curriculum materials but were not generally part of the usual instructional process" (1972, p.335).

The remaining reviewers are more critical. Lawlor (1972) criticises the lack of meaningful validity information though he states that the items "appear to be well selected from a survey of curriculum material".

Noll (1972) considers that "it is in this matter of validity that the Boehm Test is most seriously open to question...". Firstly, "no information is given as to what materials the author examined and only vague statements descriptive of methods used to determine what concepts were to be retained as a basis for the test, ... Such statements as 'considerable frequency', 'seldom if ever', 'adequate transitions', and 'relatively abstract' are too general and imprecise to convey any adequate understanding of the criteria" (ibid., p.335). Secondly, Noll notes that Boehm "supplies no information on the question of how necessary these concepts really are to 'achievement in the first years of school.'" (ibid., p.335) He queries the assumption that there is "any relation between mastery of them and quality of school achievement".

Freeman (1972) also deplores the lack of validity evidence and, additionally, points out that even if "the selected concepts are valid as to curriculum content, they cannot represent the entire range of concepts that a child should know, ideally, or needs to know" (ibid., p.337). He is critical of Boehm's lack of consideration of recent research and publications on individual differences and the methods they demand and of her ignoring of "the fact that children's differences in concept mastery have for many years been tested, in one way or another, by means of individual and group tests of intelligence, although these were not devised exclusively for that purpose..." (ibid., p.337).

Proger (1972) severely criticises the BTBC "on the grounds of insufficient validity studies" and the neglect of predictive, construct and concurrent validity.

Dahl (1973) argues that "reliability and validity estimates seem inappropriate for the stated purpose of the test." (p.65)

"... the purpose of the test is to diagnose attainment of specific concepts. Therefore, it should be looked at as 50 separate tests, each measuring attainment of a single concept. The split-half reliability coefficients reported in the Manual are therefore not relevant to the criterion-referenced notion of 50 tests. If one does consider the BTBC as representing 50 tests, the reliability must be seriously questioned, as there is only 1 test item per concept. Certainly it is risky at best to assert that a child has or has not formed a given concept on the basis of a single test item.

Validity of the BTBC suffers from an incompleteness similar to that of the reliability. The domain of examples of situations illustrating a given concept cannot be properly represented by a single test item. However, this shortcoming might be overcome by some empirical evidence. If the author could demonstrate that successful performance on an item is correlated highly with successful performance on curriculum tasks related to that item, predictive validity could be claimed." (Dahl, 1973 p.64)

The internal reliability of the BTBC is deemed good by Noll and by Proger and fairly good by Freeman (all op. cit.).

Dahl discusses the ambivalence concerning the BTBC as a norm-referenced or a criterion-referenced test.

"On one hand, the author states that the primary purpose of the test is such that the use of norms is not indicated. On the other hand, norms are provided to the teacher but the author admits they are not representative of the nation as a whole" (ibid., p.64).

Also Freeman, McCandless and Proger all draw attention to limitations in the standardisation sample though McCandless recognises wide sampling in "reasonably representative school systems." (all op. cit.) Lawlor considers the standardisation sample adequate but regrets the absence of pre-kindergarten subjects.

Three reviewers, McCandless, Noll and Lawlor, criticise the BTBC as too easy for first grade high socio-economic groups and for grade 2 pupils of any social class.

Proger notes that the "consistency of scores over time cannot be gauged from the manual, since test-retest coefficients are not provided" (op. cit., p.336). However, he commends Boehm for "her discussion of confidence intervals and standard errors of measurement" and remarks that in "the areas of conceptualization, standardization, interpretation, reliability etc., the Boehm test is better than most new test instruments appearing on the market. However, this is not saying much!" (op. cit., p.336).

The BTBC booklets are considered to be of high quality (Noll; Proger), well designed (Dahl) and "are easily handled by young children" (Lawlor). The drawings are seen as clear and concise (Lawlor; Smock), "generally clear though a few seem ambiguous" (McCandless). Both Lawlor and McCandless note the use of non-white subjects in the illustrations. On the other hand, Dahl considers that the "quality of the art work is generally mediocre but adequate. However some of the pictures are poor enough to be misleading for the examinee" (op. cit., p.63).

The BTBC is praised for its straight-forward directions by Dahl, Lawlor, Noll and Smock and for its administrative simplicity by Dahl, Noll, Proger and Smock. McCandless "is skeptical that it can be feasibly administered to groups of (kindergarten children) without a generous supply of proctors" (op. cit., p.335). Dahl remarks that examinees "seem to enjoy taking the test".

The scoring and recording of the BTBC are regarded by Dahl and by McCandless as clear and easily accomplished. "Appropriate qualifications and cautions are presented in the manual regarding interpretation of obtained scores" (Smock, p.335).

"Interpretation of test results is clearly and carefully delineated in the manual" (Proger, p.335). Lawlor also comments that the tables are easy to interpret. However, he "found no statement for the teacher regarding the cut-off level determining the difficulty of a given item for a grade level" (op. cit., p.337). Dahl returns to the fundamental problem described earlier: "Interpretation of the results is somewhat unclear, since the author seems to make contradictory statements relating to the purpose of the test" (Dahl, 1973, p.65).

McCandless regards the final item selection as having been "made according to conservative and acceptable principles". Smock notes " a fairly even increase in percent passing across the age levels". Lawlor is critical of the "preponderance of items dealing with space" and of "overlapping in that 7% of the items test two concepts and 6% tap a triad in one item" (op. cit., p.336).

Dahl has reservations about the test's intended purpose (op. cit.) but the remaining six reviewers accept the usefulness of the BTBC as a screening and diagnostic device and as a guide for instruction, remediation and evaluation whilst Proger notes that it "has implications for the disadvantaged and the handicapped." However, qualifications are expressed:
"Limitations ... should make the user cautious in its deployment." (Proger, p.336)
"This type of instrument can be useful only in providing teachers with symptoms, or clues, as to children's concept mastery or ignorance." (Freeman, p.337)

"Used thoughtfully, it can be quite useful to teachers." (McCandless, p.335) If there is no relation established between mastery of the Boehm concepts and achievement in the first years of school then

"A teacher who gave the test, tabulated the results, and proceeded with careful remedial instruction, would seem to be wasting time." (Noll, p.p. 335/6)

The BTBC manual's suggestions for remedial instruction emphasise that training on the test items is to be avoided. Noll sees some difficulties with this

"... if, as appears likely, synonyms for such concepts as top, bottom, whole, every, etc. are to be found and used equating them directly with their counterparts in the test." (ibid., p.335). It is doubtful that teachers would share Noll's interpretation of Boehm's warning. She does not specify that the actual words which label the concepts are to be avoided.

Boehm's own summary of reviews of the BTBC is as follows:

"Overall the strengths of the BTBC according to the reviewers include that it is well-designed with straightforward directions, usually clear pictures, reasonably representative norms, and a good discussion of results which teachers should find useful for instruction. The major weaknesses of the test were seen as the incompleteness of the validity, documentation of frequency of concept use in curricular materials, and evidence regarding the importance of the 50 concepts assessed to achievement. Another source of comment suggested that the basic intent of the test relates to criterion-referenced assessment rather that norm-referenced assessment, making much of the statistical data presented in the manual inappropriate.

A study recently completed by Boehm, Kaplan and Preddy (1979) deals with the issues of content validity and needs to be incorporated into a revised manual." (Boehm, 1980 p.1)

This last named study is reviewed in section D of this chapter together with other validity studies.

C <u>Bias and the BTBC</u>

Silverstein, Belger and Morita (1982) applied one of Jensen's (1980) internal criteria to the BTBC standardisation data. Boehm (1971) reports the percentages of children in each socio-economic group of the standardisation sample who passed each of the 50 items of the BTBC. In an unbiased test, the rank orders of item difficulties should not differ markedly between groups (Jensen, 1980).

"Spearman rank correlation coefficients were calculated between the item difficulties for all the groups in both samples. Overall, the median correlation for children of different socioeconomic levels (.87) proved to be identical to the median correlation for children of the same socioeconomic level. Various finer grained comparisons also were made, with similar results. For example, the median correlation for children of the same socioeconomic level tested at the beginning of the school year and at the midyear was .93. The median correlation for children of different socioeconomic levels tested at the same time of the school year was .91. The median correlation for children of different socioeconomic levels, lagged so as to allow for the differences in mean scores noted previously, was .92." (Silverstein et al., 1982 p.32).

Thus by the criterion applied by these authors "there is little evidence of bias in the BTBC." Items that are difficult for one group were found to be difficult for all groups. "... this suggests that children of different socioeconomic levels master basic concepts in about the same temporal order, differing only in the rate at which they do so" (ibid., p.432).

The BTBC is normed on an urban standardisation sample. These urban derived reliabilities were compared with data from rural children (Houck, Biskin and Regetz, 1973). The BTBC was administered to all 121 kindergarten and 293 first grade pupils in a rural Appalachian county. The criteria for socio-economic status which Boehm (1971) had employed were not available (ibid.). The rural children were categorised according to fathers' occupations and levels of education. There were 162 lower-, 242 middle- and, in Grade 1 only, 10 upper-class children. The reliability estimates were computed by a split-half reliability coefficient as with Boehm's sample.

The reliabilities of scores for the rural kindergarten sample are .88 for lower- and .52 for middle-class whilst Boehm's corresponding figures for her urban sample are .86 and .90. First grade reliability coefficients are .83, .71 and .42 for rural lower-, middle- and upper-classes whilst the corresponding urban figures are .82, .82 and .76. Differences between the urban and rural middle-class reliabilities are significant at the .01 level for both kindergarten and first grade. The even greater absolute difference between the urban and rural upper-class reliabilities is not statistically significant (rural upper-class N=10). Standard deviations of scores of urban kindergarteners were 8.1 (lower-) and 8.0 (middle-class) as compared with 7.6 and 4.7 for rural children. The corresponding deviations for first-graders were 5.5, 4.5 and 3.7 (upper-class) for the urban sample and 6.6, 3.8, and 3.0 for the rural sample. In all cases except one (lower-class, Grade 1) there was a greater variation in the urban sample (ibid.). Despite the likelihood of lower reliability estimates with small samples Houck et al. warn that "because of the large discrepancies in reliability estimates caution must be exercised when this test is used to assess rural children" (ibid., p.432).

Differences in mean scores and in intercorrelations between a group of preschool tests, which included the BTBC, and demographic variables are observed by Ernhart, Spaner and Jordan (1977). This investigation of the validity of preschool screening tests is reviewed below in Section D. Data from 97 black and 188 white children are treated separately "because the races differed on a number of social indicators" (ibid., p.79). So although the mean scores of the black children were lower than those of the white on seven of the nine tests, no conclusions about racial bias in the BTBC can be drawn from this investigation. Cole (1981) notes that understanding item bias is complicated by, among other things, the lack of homogeneity of groups labelled 'black' or 'white'. (The only measures which showed slightly higher mean scores for 3 year old black children were the Ambulation and Manipulative Subscales of the Preschool Attainment Record. The BTBC mean scores of 4 year olds were 15.48 (SD 4.67) for white and 11.57 (SD 4.93) for black children.)

Baglin (1981) postulates a "self-selection" bias in many achievement test norming samples. Through the analysis of information from the three largest national testing companies in America (The Psychological Corporation, CTB/McGraw Hill and Houghton Mifflin/The Riveinde Publishing Company) Baglin finds that the school districts actually used in three major norming studies "constitute a very different set of districts from the original random samples drawn..." (ibid., p.97). "Only about one out of five invited districts agreed to participate, leaving ample room for self-selection biases to creep in" (ibid., p.98). The tests under investigation were the Metropolitan Achievements Tests, 1978 (MAT 78), the California Achievement Test 1977 (CAT 77) and the Iowa Tests of Basic Skills/Achievement and Proficiency, 1978 (ITBS/TAP 78).

Additional evidence from questionnaires sent to school districts strongly supports the hypotheses:

"A district's likelihood of agreeing to participate in a national norming study of a standardized achievement test is dependent upon that district's
(1) use of (a) that same achievement test series in an earlier edition (<u>same test</u>), or (b) any achievement test series from the same publisher (<u>same test publisher</u>) ;
(2) use of instructional materials from that same test publisher or its parent company (<u>same instructional materials publisher</u>)."
(ibid., p.97).

Districts may choose a specific test series for regular use "because they will be able to score higher, since it more closely fits their curriculum" (ibid., p.107). Baglin concludes that "the three purportedly random "national" norming samples are not at all random... ...each test is normed on a nonrepresentative, nonrandom segment or cross section of the population. What is occuring is a form of psychometric inbreeding" (ibid., p.104).

The self-selection identified in Baglin's study
"results in school districts with certain characteristics being overly

represented. This leaves a potential that they and similar districts are likely to improve their schools' scoring performance on the test in question. Since this situation exists in each of the three major test publishers' most recent norming studies, it is not unreasonable to suspect that it is true of recent norming studies in general" (ibid., p.104).

The standardisation sample for the BTBC was drawn from sixteen cities for 'beginning-of-year' norms and from five cities for 'mid-year' norms (Boehm, 1971). Boehm considered that it was "unnecessary to select standardization samples representative of chidlren ... in the nation as a whole" (ibid., p.19).

"School officials in each cooperating city were asked to provide classroom groups from schools with a fairly wide range of socioeconomic background. ... No formal specifications for this selection were given; the choice was left to the judgment of the administrative personnel" (ibid., p.19).

In the light of Baglin's (1981) findings one wonders about the 'self-selection bias' which might have been operating.

Three statistical aspects of bias were examined for use of the BTBC with white and Mexican-American children: consistency of reliability across groups, cross-group factorial congruence, and equivalence of regression systems in the prediction of school achievement (Reynolds and Piersel, 1983). No empirical evidence of bias could be determined (ibid.).

D Validity of the BTBC

"A test's validity is the extent to which scientifically valuable or practically useful inferences can be drawn from the scores" (Jensen, 1980 p.297). The demonstration of a test's validity may be based on any one or any combination of the four main types viz. content validity, concurrent validity, predictive or criterion validity and construct validity (ibid.). Generally speaking the studies to be reviewed below will follow the order just mentioned but several are combinations of the standard types of validity.

"A test has content validity to the extent that the items in the test are judged to constitute a representative sample of some clearly specified universe of knowledge or skills" (ibid., p.297).

Boehm, Kaplan and Preddy (1980) examined (a) the extent to which the 50 basic concept terms measured by the BTBC occur in printed materials, (b) how extensively these 50 terms appear in specific reading and mathematics instructional materials the child is likely to receive in KG, grade one and grade two, and (c) the frequency with which these terms occur in verbal teacher instructions.

They report that 41 of the BTBC terms, in one of their comparative forms, appear among the first thousand most frequently used words on the Thorndike-Lorge list whilst two additional concepts were represented by a synonym or antonym ('alike' by 'same'; 'separated' by 'together'). Seven of the 50 concept terms (equal, inside, medium-sized, matches, row, skip, zero) did not appear amongst these first thousand words. Boehm counteracts criticisms of the datedness of the Thorndike-Lorge list (1944) by citing studies undertaken in the late sixties which showed all of the BTBC terms appearing in grades K to 3 reading materials of nine publishers except for 'least' and 'equal' which appeared at grade 4, 'separated' at grade 5 and 'medium-sized' at grade 6. (ibid.).

Boehm et al. surveyed five sets of reading and five sets of mathematical materials chosen from the most commonly used curricula material in the USA. All 50 BTBC terms were employed: 9.5% and 8.8% of all words on sampled pages, of reading and mathematics work books respectively, were BTBC terms. Boehm's paper provides no details of the frequency of repetition of each of the concept terms.

Boehm et al. assert that one out of every two verbal directions contain a direct BTBC term or its synonym, antonym or comparative form. "Directions" are defined as "verbal interactions, in command form, which describe how a task is to be performed and/or requests performance of a behavior or behaviors" (ibid., p.6). In 711 verbal directions by nine teachers (three in each of grades K, 1 and 2) 33 of the 50 BTBC terms and 18 other words (antonyms, synonyms or comparative forms) were used (ibid.).

Kaplan devised a task, "The Directions Game", in order to study the extent to which children are able to follow teacher directions (ibid.). There was a strong relationship (r = + .71) between concept mastery, as measured by the BTBC, and performance on the Directions Game. "Kaplan concluded that, while the complexity of directions does not appear to change across grades, children's ability to execute directions correctly does change, increasing across grades, especially in grades K - 2" (ibid., p.7). The second of Kaplan's conclusions is hardly surprising but it seems more likely that primary teachers, in the UK at any rate, <u>do</u> increase the complexity of their directions across six age-groups. Personal observation of reception class teachers and teachers of ten year olds does not suggest congruence of complexity in their verbal directions. However, without evidence, this view remains conjectural.

Kaufman (1978) used the content of the BTBC in an analysis of the test directions of the WPPSI, the ITPA, the Stanford-Binet and the McCarthy Scales. He asked: "To what degree do the directions spoken by the examiner assume the young child's knowledge of basic concepts?" (ibid., p.207) The

BTBC basic concepts assumed to be known by the young child are listed below for the four tests.

ITPA	McCarthy (General Cognitive Scale)	Stanford-Binet (Years II-VII)	WPPSI
No concepts are assumed	after	alike	after
	alike (same)	around	alike (same)
	away	difference	different
	different	other	farther
	other	whole	half
	over		in order
	whole		inside
			middle
			next
			other
			row
			side
			skipping
			top

(ibid., p.208).

Kaufman considers that lack of knowledge of these basic concepts may depress scores on the abilities presumed to be measured by the various tests especially on the WPPSI.

Kaufman also points out that

"it is ironical that even the directions for administering the Boehm Test of Basic Concepts include two Boehm concepts ('different', 'around'), one which is the opposite of a Boehm concept ('before') and five which are on the supplementary list of easier concepts ('line', 'up', 'change', 'down', 'open'). The inclusion of basic concepts in the directions for the Boehm test itself should not affect the child's performance, since the examiner must make sure that the child understands the task at hand before proceeding." (ibid., p.209).

Meissner (1975) reports that inner city children who responded to the BTBC items correctly did not necessarily use these terms in their everyday language. The subjects were black children in second and fourth grades in South Bronx. Ten of the BTBC items were administered as a comprehension task and the same ten concepts were used in a child pairs communication task with toy objects.

"Even though there were instances of Black English grammatical forms in the speech of some speakers, there were very few novel expressions of the concepts. Most of the speakers who used the concepts used the same terms for these concepts as used in the Boehm Test (1969). The two major exceptions were the speakers' more frequent use of 'middle-sized' instead of

Boehm's 'medium-sized', and 'middle' instead of 'center'. 'In back of' and 'behind' were used about equally often, as were 'in a line' and 'in a row' (ibid., p.26).

The remaining six BTBC items were 'at the top', 'through', 'next to', 'second', 'most' and 'half'. The major findings from this study highlight the inadequate listening and questioning skills of these children.

The "term concurrent validity should be used only to refer to ... the correlation of a previously unvalidated test with an already validated test" (Jensen, 1980 , p.302). Concurrent validity evidence is reported on the Central Susquehanna (CSBT) modification of the BTBC (Booklet 1. Form A) (Levin, Henderson, Levin and Hoffer, 1975). Local Appalachian words were substituted for certain nouns in the instructions (e.g. 'couch' for 'sofa'). The normative sample is "narrowly defined ... as white, disadvantaged and preschool" (ibid., p.136). Subjects were 166 children enrolled in either a rural Headstart program or a Title IV-A centre in a small industrial city; ages ranged from 2:10 to 4:3. No significant differences were found between the two cohorts nor between the sexes.

Correlations of .62 with the Peabody Picture Vocabulary Test (PPVT) and .41 with the California Preschool Social Competency Scale (CPSCS) were obtained but when age was partialled out the resulting coefficients were .57 and .29. No information is given on the number of subjects tested with these two standardised tests. Levin et al. claim that their validity data "support the idea that the test does measure knowledge of basic concepts, though it was beyond the scope of (this) research to test whether these concepts indeed are critical in elementary school" (ibid., p.133).

> "A series of studies have indicated that when children receive individual tutoring aimed at teaching them the concepts measured by the CSBT, their scores on the CSBT go up. Typically, tutored bhildren not only show a statistically significant increase in CSBT scores, but catch up with both age norms and untutored children with higher initial scores on the CSBT, PPVT, and CPSCS. This sensitivity of the CSBT to intervention also supports our construct of what the test measures. Better controlled studies and more detailed data analyses are needed here, however, before this evidence on change can be properly evaluated" (ibid., p.138)

No details are reported of these intervention studies. Indeed, it may well be that such factors as motivation, increased self-confidence and test sophistication are operating as noted in the previous chapter.

Ernhart, Spaner and Jordan (1977) examined the validity of preschool screening tests some of which were administered at 3 years of age and some at 4 to 188 white children and 97 black children. The tests were the PPTV, three subscales of the Preschool Attainment Record (PAR), the Verbal Language Development Scale (VLDS), the Copy-Forms Test, the Quick Test (QT), the Preschool Inventory (PSI) and individual administration of Booklet 1 of the BTBC.

The BTBC scores at 4 years correlate with demographic variables and the other measures as follows:

For white children (N=188 except for Copy-Forms N=169, VLDS N=65, QT N=123) SES -.33, maternal education .34, tests previously administered at 3 years .12 (PAR Ambulatory), .19 (PAR Manipulative), .27 (PAR Communication), .53 (PPVT IQ) and tests at 4 years .68 (PSI), .50 (Copy-Forms), .30 (VLDS) and .63 (QT IQ).

For black children (N=97 except for Copy-Forms N=56 and the VLDS and QT N=18) the corresponding coefficients are -.10, .14, .39, .26, -.08, .49, .66, .37, .28, .52. The mean BTBC score for white children was 15.48 (SD 4.67) and for black children 11.57 (SD 4.93) (ibid., from Tables 1 and 2 p.p.82/83).

"Among the tests evaluated in this study, those which were most characteristic of traditional intelligence tests show the greatest concurrent and predictive validity. Furthermore, they are associated with those demographic variables which are usually related to intelligence test scores. These tests are the Peabody Picture Vocabulary Test, the Preschool Inventory, the modified Boehm Test of Basic Concepts, and, at a predictably lower level, the Copy-Forms Test. The first three appear to be reasonably valid brief tests appropriate for screening and for program evaluation..." (ibid., p.87).

Thus the investigation by Ernhart et al. seems to show that the PPVT may, to some extent, predict performance on the BTBC and that performance on the BTBC correlates fairly well with current performance on the PSI, the QT(IQ) and the Copy-Forms Test but only weakly with the VLDS. There is no evidence from this investigation of the predictive validity of the BTBC itself.

A highly significant correlation (.71) was observed between the BTBC raw scores and the ITPA mean scaled scores of 107 children (Smith, 1977). Each test was administered early in the first term of infant school; all the children had experienced a full year of half-day nursery class; ages ranged from 4:1 to 5:3; almost all the children had working-class backgrounds and many were socially disadvantaged[1]. The BTBC mean raw score was 25.99 and the standard deviations for each of six schools ranged from 6.5 to 11.1. The ITPA mean scaled score was 35.75 (SD 5.3) and the means for each of eight sub-tests ranged from 32.75 (Verbal Association; SD 6.2) to 40.65 (Verbal Expression; SD 9.7). (The standardised ITPA mean for each sub-test is 36 with a SD of 6.)

This study also provides some evidence of the relationships between a number of variables and subsequent performance on the BTBC. A small but highly significant correlation (.39) was observed between the BTBC raw scores of 62 children and raw scores on Binet non-verbal items some two to seven months earlier at age 4:0 - 4:3. Gains made by 107 nursery children on the ITPA sub-test Auditory Association, between the pre-test at the start of the nursery year and the post-test at the start of the following year, correlated .36 ($P < .01$) with their BTBC raw scores several weeks after the post-test ITPA. The correlation of the same children's record of attendance during their year in the nursery class and their subsequent BTBC raw scores was also highly significant though relatively small (.28). A correlation of .26 ($P < .05$) was found between subsequent BTBC raw scores and the mean length of utterance (MLU) observed during the previous term. (A child's MLU was calculated from recordings of conversations between children and with adults.)

1 See page 8.

No significant relationships were detected between the BTBC scores and any of the following variables: gains or losses (betwen pre- and post-test) on ITPA mean scaled scores and on each of six sub-tests (Auditory Reception, Auditory Memory, Visual Reception, Visual Association, Verbal Expression, Grammatic Closure); language complexity scores (noun phrase + verb phrase); the number of sustained language interactions between child and teacher; the number of interactions between child and teacher.

"The only ITPA sub-test where the gains made by the children correlate significantly with Boehm scores is Auditory Association. It is probable that skill in listening and attending enables a child to score well in both tests. That high scorers on Boehm tend to make greater gains on AA, during their nursery year, suggests that the nursery may have helped to develop these skills. The significant correlation between attendance and AA gains supports this." (Smith, 1977 p.150)

"The low correlation of gains made on the ITPA MSS with Language Complexity, MLU and Boehm scores were not unexpected: the bottom third[1] made the largest gains but would be unlikely to score high on the other three measures" (ibid., p.150).

Piersel and McAndrews (1982) examined the relations of scores on the BTBC to tasks such as Draw-a-Child and knowledge of colours, shapes, numbers, letter recognition and counting skills. Complete data were obtained for 123 kindergarten children in a semi-rural district whose mean age at initial testing was 5:3 and who were primarily of lower-middle and lower socio-economic status. The children were tested during the week prior to entering kindergarten and at the end of kindergarten. In addition to the BTBC the initial screening tests included tests 12 and 13 of the McCarthy Scales (Draw-a-Design and Draw-a-Child) and a locally designed criterion-referenced test of basic skills which "included knowledge of 8 colors, 6 shapes, recognition of numerals, identification of 26 letters, counting objects, writing game, and oral-rote counting" (ibid., p.784). The criterion was the SRA, 1978 administered 15 months subsequent to the initial testing.

[1] The sample was divided into thirds on the basis of their pre-test ITPA mean scaled scores.

This has

"5 basic subtests designed to measure visual discrimination, auditory discrimination, letter recognition, direction following-listening, and basic mathematics skills. This criterion was selected by the school district personnel as the standardised achievement test which most closely measured what was taught during the elementary grades" (ibid., p.784).

The lowest concurrent relationships between the BTBC and other tests were found for Draw-a-Design, Draw-a-Child and knowledge of six basic shapes (rs = .26, .21, .24 respectively; $p < .01$). "The Boehm test has moderately strong concurrent relationships to knowledge of colors, counting skills, number recognition, and letter recognition (rs = .47, .51, .48, .37, .43, respectively" $p < .01$) (ibid., p.785).

"In terms of predicting performance at the end of kindergarten, the Boehm scores demonstrate similar correlations with these measures even though these tests were administered 8 mo. later" (ibid., p.785). The product moment correlations of the BTBC and the SRA Achievement Series administered 15 months later are reported as .54 (total SRA), .56 (total mathematics), .44 (listening), .43 (total reading), .30 (both auditory discrimination and letters), .22 (visual discrimination). All these correlations are significant ($p < .01$).

"The SRA listening subtest and the Boehm score suggest a common requirement of following directions. This hypothesis is further supported by the findings of Piersel and Reynolds (1981) that the Boehm test is essentially a one-factor test of ability to follow directions" (ibid., p.786. See also Chapter 4 and below.)

The report of Piersel and McAndrews study does not specify individual administration of the BTBC so presumably, like the SRA, it was group administered.

"Generally, the present results support the use of the Boehm Test of Basic Concepts to predict current as well as future academic achievement at Grade 1" (ibid., p.786).

"One consistent finding in this study and previous studies (Estes et al., 1976; Busch, 1980; Steinbauer and Heller, 1978) is the moderately high correlation with counting and arithmetic skills. While no one has speculated on this, inspection of the Boehm test and SRA Arithmetic subtest yields several items assessing quantitative thinking, sequencing skills, as well as basic numerical concepts" (ibid., p.786).

A predictive validity study by Estes, Harris, Moers and Wodrich (1976) supports Boehm's assertion that mastery of the BTBC concepts is related to achievement in the first years at school. First grade pupils were tested with the BTBC in September and with the Primary I Battery of the Stanford Achievement Test (SAT) in the following May. The median correlation between SAT subtest and total BTBC scores was .47 (N 278 $p < .01$). The largest intercorrelation was for SAT Arithmetic and the BTBC items categorised as "Quantity" ($r = .53$) whilst SAT Spelling subtest exhibited the lowest correlation with each BTBC category. The schools used were all in one metropolitan area containing pupils from lower-middle to middle level of socio-economic status. 75% of the subjects answered correctly all 50 of the BTBC items. Estes et al. predict that higher correlations between BTBC scores and SAT scores would emerge among kindergarten and disadvantaged pupils because the variances of the BTBC scores would increase (ibid.).

The predictive study by Steinbauer and Heller (1978) did include kindergarten children but the sample was taken from just one "upper-middle to upper-class school". 94 second-and third-graders who had taken the BTBC in the kindergarten were given the SAT, form W. Steinbauer and Heller report high positive correlations between the BTBC and Spelling (.54 in grade 1 and .72 in grade 2), and Language (.53 and .93). Arithmetic Computation and Arithmetic Concepts correlated positively with early BTBC scores (.58 and .41 for Computation and .65 and .47 for Concepts). It is unclear whether or not the teachers knew the results of the early BTBC testing. If they did, and acted on them remedially, then either their effectiveness as teachers is questionable or this is an instance of pupils fulfilling teachers' expectations. (If it were argued that defective concept mastery is not remediable, however able the teacher, this would refute the declared purpose of the BTBC.) If some of the children raised their level of achievement through remediation efforts then the correlation with later tests would be reduced.

On the other hand, as the mean scores and deviations are not reported it is unclear what sort of picture the obtained coefficients reflect. This uncertainty endorses the plea made by Hunt and Lansman (op. cit.) to see to what extent trends are typical of individual subjects.

If the teachers in this study did not know the results of the BTBC testing then the claims made by Steinbauer and Heller that the BTBC is successful in predicting school achievement would be acceptable for this particular type of population i.e. children of high socio-economic status.

Moers and Harris (1978) examined the postulate that remediation of the conceptual deficits detected by the BTBC might enhance school achievement. All first graders attending a school in a predominantly lower-middle to middle socio-economic level area were given the BTBC (Form A) early in September. Complete data were collected for 54 of the 66 children scoring below the 40th percentile (local norms). These pupils were randomly assigned to one of three treatment groups (experimental, placebo and control). The experimental treatment was designed to conform to Boehm's suggestions for remediation as expressed in the test manual. At mid-year Form B of the BTBC was administered and no special instruction took place after this. Form A was readministered at the end of the school year as was the Stanford Achievement Test (Form W). No significant treatment effect was detectable in the BTBC results. "Post hoc tests indicated that the three groups increased their scores from pre- to posttest and from posttest to delayed posttest" (ibid., p.85).

A significant treatment effect ($p < .06$) was revealed by a one-way multivariate analysis of covariance for the six SAT subtests and the BTBC delayed posttest considered simultaneously, with the pretest treated as covariate. "SAT Spelling, Word Meaning, Word Study Skills, and Paragraph Meaning subtest scores were most strongly related to treatment group membership..." (ibid., p.85). The mean grade equivalent on SAT subtests was 1.71 for the experimental group, 1.67 for the control group and 1.61 for the placebo group.

"Assuming replication, whether an effect of this magnitude justifies 600 minutes of small-group instruction is debatable. Of course, other BTBC-based instructional programs might well result in a more pronounced treatment effect. Additionally, disadvantaged or younger children who have mastered fewer basic concepts prior to formal instruction might be expected to profit more clearly from the training program employed in this study" (ibid., p.86).

Moers and Harris advance two speculative explanations for the lack of treatment effect observed in the BTBC results. "First, it may be that the BTBC is not sufficiently sensitive to reflect important differences that really existed" and second, it is possible that "experimental subjects learned to be better learners and/or performers on structured tasks" (ibid., p.86). Literature on this last possibility has been cited in Chapter 4.

The BTBC was employed in a study designed "to determine the best combination of tests or subtests in a research battery which, when administered to beginning first-grade students, would result in the most efficient prediction of reading achievement" (Busch, 1980 p.39). The results from a battery of seven tests administered to 1,052 children, representative of urban, suburban and rural populations across the State of Missouri, were compared with end-of-year first grade reading achievement as measured by the Gates-MacGinite Reading tests. The data were analysed by correlational and regression analyses.

Ages ranged from 61 to 96 months with a mean age of 79.24 months (SD 4.40). The BTBC scores ranged from 16 to 50 with a mean score of 43.49 (SD 4.63). The BTBC correlated .56 with the Gates test and with the Behavior Rating Scale, .46 with Visual Discrimination of Letter Forms (Pre-Reading Screening Procedures, PRE-R 1) .43 with Visual perception-memory-motor (PRE-R 5), .42 with Visual Discrimination of Words (PRE-R 2), .41 with Copying (PRE-R 4), .39 with Visual Perception Memory (PRE-R 3), .36 with Visual-Motor Integration (VMI), .35 with Letter Knowledge (PRE-R 7) and .28 with Auditory Discrimination (PRE-R 6), ($p >$ .01 in each case). Correlations between the other measures and the Gates test ranged from .68 (SESET

Letters and Sounds) to .23 (PRE-R 6). By use of stepwise regression Busch finds the SESET subtest Letters and Sounds making the greatest contribution to the dependent variable. Next in order are IQ and the Behavior Rating Scale.

Busch concludes that "it is possible to predict with a substantial degree of accuracy the reading achievement of first-grade students" (ibid., p.47). He asserts that this can best be done by a combination of the assessment of the ability to recognise upper and lower case letters and beginning sounds, a group test of intelligence and the rating of children's behaviour by teachers.

"Construct validity becomes a consideration as soon as we have some theory (or 'construct') as to the psychological nature of the trait that we wish to measure" (Jensen, 1980 p.303). "A test devised to measure the trait is said to show construct validity if the test predicts the behavior in specific situations that would be deduced from our theory of the trait. ... The task of construct validation is never really completed. A test's construct validity is further enhanced by every ... theoretical prediction that is borne out in fact" (ibid., p.403).

In order to "identify some of the convergent and discriminant characteristics of the BTBC a multitrait-multimethod matrix was used (Campbell and Fiske, 1959). The two 'Methods' were the areas of verbal and perceptual-motor development. The 'Traits' were comprehension, relational thinking, spatial thinking, and general intelligence" (Brown, 1975 p.1). This method permits the examination of interrelationships among tests thought to tap a single trait by different methods together with the interrelationships among a variety of traits (ibid.).

It was predicted that the BTBC is "primarily a measure of spatial abilities strongly related to verbal development as measured by 'traits' in the 'Verbal Method' and possibly, though less strongly, related to visual motor development as measured by 'traits' in the 'Perceptual-Motor Method'" (ibid., p.2).

The measuring instruments used by Brown were Booklet 1 of the BTBC, the Miller-Yoder Test of Grammatical Comprehension, the Peabody Picture Vocabulary Test (PPVT), the Development Test of Visual Motor Integration (VMI), the Draw-a-Person Test (DAP), the Picture Association subtest from the Hiskey-Nebraska Test of Learning Aptitude, letter recognition and classification tasks. The latter were designed to tap the child's ability to classify at the perceptual level (similarity of colour, shape or number), the functional (use e.g. 'you eat it') or superordinate levels (class labels). As they were designed specifically for this project no reliability or validity information is available (ibid.).

The subjects were mainly white from a small mid-west farming community. Almost all the children who expected to enter kindergarten were tested in groups of 8 (N = 216); ages ranged from 4:3 to 5:11.

It was hypothesised that success on the BTBC needs verbal comprehension of words and perceptual ability in the Piagetian sense i.e. topological, projective and Euclidean relations developing in parallel. The BTBC items (Book 1) classified as topological were: through, away from, next to, inside, over, between, nearest, behind. Those classified as projective were top, farthest, around, widest, row, behind, second. Those classified as Euclidean were corner, whole, middle.

Brown finds the BTBC's strongest relationship to be with the VMI test. "The relatively strong relationship of the BTBC to all the other tasks suggests that it reflects both perceptual-motor and verbal development" (ibid., p.74).

"When one examines what is related to what it becomes evident that 'verbal' and 'perceptual-motor' were not clear-cut divisions as operationalized in this study. The correlational values ranged from .119 to .487 over 28 pairs... The PPVT shared the least commonality with other tasks, appearing 5 times in weaker 8 correlations. Among the strongest relationships was the BTBC appearing 4 times in the last 12 pairs" (ibid., p.76).

All tasks correlated positively with SES as determined by parental education and the father's occupation. The coefficients ranged from .181

(VMI and occupation of father) to .389 (PPVT and education of father). Most were in the range .2 to .3+. Brown suggests these values might have been higher had the SES range been broader.

Brown acknowledges the limitation imposed by the narrow range of socio-economic levels in her sample. She also suggests that more discriminant variables were needed and that an "item analysis of the BTBC might shed more light on some of the present correlations" (ibid., p.85).

Steinert (1978) examined the construct and criterion-reference validities of the BTBC. She projected that performance on the BTBC is representative of the child's developmental level of language and of his level of concept maturity. Construct validity was inferred by examining the relationships of measures of language development and cognitive development with the BTBC. Language development was measured by the Carrow Test of Auditory Comprehension of Language (TACL) and the Carrow Elicited Language Inventory (CELI). Cognitive development was assessed by performance on Piagetian type tasks of classification, seriation and number conservation.

Predictive validity was examined by administering the Metropolitan Readiness Tests (MRT) in the spring following the autumn testing with the BTBC, and by end-of-year placements of pupils.

The subjects were 144 kindergarten children in one school district which was predominantly middle-class rural/suburban. Ages ranged from 5:1 to 6 years (mean age 66.326 months). The kindergarten curricula was similar for all schools in the district.

The canonical correlation of the Piagetian and language predictors and the BTBC criterion was .861. "R^2 = .741 and represents the total proportion of variance redundant between the Piagetian tasks and Carrow tests and the BTBC on the trait measured by canonical variate <u>one</u>" (ibid., p.93).

The "largest coefficients were BTBC Space (.961) and TACL Morphology (.856) with canonical variate 1. ... This suggests that BTBC performance corresponds to the ability to interpret morphological features in language more than the ability to interpret single vocabulary items (vocabulary subtest) or more complex syntactical structures (syntax subtest)" (ibid., p.112).

Minimal commonalities between the BTBC and the CELI indicate that expressive or imitative language abilities are not strongly related to performance on the BTBC (ibid.).

By knowing a subject's Piagetian and Carrow scores approximately 19.5% of his performance on the BTBC can be estimated; by knowing a subject's scores on the BTBC approximately 46.1% of his performance on Piagetian and Carrow tests can be estimated (ibid.).

A multitrait-multimethod matrix shows the relationship between the BTBC subtests and Piagetian tasks. (Steinert had reclassified Boehm's item categories of space, time, quantity and miscellaneous into five categories n. space, number, seriation, time and classification.) The BTBC subtests correlations are all significant ($p < .01$) and are moderately positive ranging from .492 to .699. The 10 Piagetian tasks correlate positively at a lower range (.097 to .414) with $p < .01$ for four of these tasks and $p < .05$ for two.

"Similarities among the rankings of correlations does not provide strong evidence for the discriminant function of the subtests on the BTBC. While a match can be found for the tests of seriation and space which seem to support the proposed relationships of cognitive concept and verbal label, actual values are only moderate (Space, $r = .512$; Seriation, $r = .439$). The correlations reveal a rather mixed pattern of relationships between the BTBC subtests and Piagetian tasks. ... While convergent validity is supported, discriminant validity is only partially upheld" (ibid., p.102).

"The strong correlations of the five subtests with the BTBC total score and the moderate correlations among the five subtests seem to suggest that the BTBC subtests are not measuring discrete abilities or knowledge. While a limited disparate function may be operating for each of these subtests, it appears that a more common underlying factor was the focus of the evaluation of construct validity in this study. The correlations with the tests of language development and cognitive development is viewed as evidence of this factor" (ibid., p.106).

Steinert suggests that the BTBC functions as "a general estimate of cognitive development. It does not appear to be the best measure of discrete conceptual clusters or cognitive structures. Piagetian tasks appear to discriminate better among the various substantive areas" (ibid., p.115).

The highest correlation coefficient between achievement test scores is that found between the BTBC and the MRT ($r = .50889$). Multiple regression analysis shows that the BTBC is able to account for the largest amount of variance on the MRT. Steinert suggests, therefore, that the BTBC measures receptive language abilities.

At the end of the kindergarten year five out of 55 children were placed in special classes or groups; four of these children's BTBC scores were in the lowest quartile. Thus the BTBC "appears able to identify students who may experience difficulty in an academic situation because of lack of knowledge of the concepts present on the test" (ibid., p.106) (Or might this be because of some other cause or multiplicity of causes?)

Steinert concludes from her study that the BTBC

"seems to function primarily as a measure of the knowledge of verbal concepts which has as a base some requisite level of cognitive development along with a knowledge of the iconic representatioon which operates in the verbal and visual stimuli present in the testing paradigm" (ibid., p.122).

She contends that her results support the original intention for the test as described by Boehm (1971). Furthermore, Steinert suggests that the power of the BTBC Space subtest (27 concept terms) may be symptomatic of a more general level of functioning required for school success. It might prove a valuable source of information on the receptive/comprehension abilities of children with expressive language difficulties (ibid.).

Piersel and Reynolds (1981) examined Boehm's (1971) classification of test items, into the four context categories of space, quantity, time and miscellaneous, from a factor analytic perspective. Data were gathered as part of a district-wide screening process. One group of 238 prospective preschool and kindergarten children (mean age 59.7 months) was tested with Form A immediately prior to the school year and a second group of 316 preschool and kindergarten children (mean age 68.0 months) was given Form B at the end of the school year. There was substantial overlap between these two groups (ibid.). Both groups were composed of white children (124 and 164) and Mexican-Americans (109 and 152). Kudor Richardson "reliability estimates were in the middle to high .80s for both groups on each form. The white means are very near to those reported by Boehm (1971) for the standardization sample, whereas the Mexican-American means are nearer the means for low socioeconomic groups reported by Boehm (1971)" (ibid., p.581).

Forms A and B were factor analysed separately. Factors were rotated orthogonally according to the varimax criterion. (See Jensen, op. cit., on the need for oblique rotation.) Although the Scree test indicated a single factor, 1-, 2-, 3- and 4- factor solutions were examined. "The 2-, 3- and 4- factor solutions failed to provide logically interpretable groupings of the items" (ibid., p.582). Neither the 3- nor 4- factor solutions "resembled Boehm's groupings of items; none of the individual factors matched any of the factors suggested by Boehm, nor did they appear to group in any logically deduced fashion. Thus, within the present group of children, the BTBC is best conceptualised as an unidimensional scale that measures a child's general level of acquisition of basic language concepts" (ibid., p.582).

Piersel and Reynolds conclude "that both forms of the BTBC are unifactorial and that they measure a consistent construct. Clearly Boehm's (1971) classification of BTBC concepts does not seem warranted" (ibid., p.583). This last statement supports Steinert (op. cit.), even though somewhat different context categories were used in each of these studies. Steinert sub-divided Boehm's category of 'Quantity' into 'number' and 'seriation' and, for example, she reclassified the item 'in order' from 'Space' to 'Seriation'.

It is doubtful that Piersel and Reynolds are justified in their claim that the BTBC is unidimensional and measures a consistent construct, namely, a level of acquisition of basic language concepts. As already discussed in Chapter 4, it is highly improbable that there is a single dimension of concept mastery for which a summary score could establish different levels.

Principal Components Analysis of 16 variables which included the BTBC produced five first order factors (Smith, 1977). Subjects were 108 nursery children and observations were made during their nursery year and at the start of the following year. Factor I appeared to be related to auditory-vocal skills. Factor II was assumed to be Vernon's g + v:ed. Five variables loaded 0.4 or greater on this factor: the BTBC (-.84), the ITPA post-test mean scaled score (-0.60), non-verbal IQ (-0.60) and nursery class attendance (-0.56).

The BTBC scores also loaded (0.70) on Factor II of the second order factors. Other variables loading .05 or greater on this factor were the ITPA post-test mean scaled scores (0.77), MLU (0.68), language complexity (0.53) and non-verbal IQ (0.51).

E The BTBC and Children with Sensory or Learning Disabilities

The BTBC was administered individually to 24 hard-of-hearing children, aged between 6 and 8 years, who spent at least half of each school day in classrooms for the normally hearing (Davis, 1974). The purpose was to investigate the performance of hard-of-hearing children, currently enrolled in public schools, on the BTBC. School records indicated that hearing loss was the only handicapping condition and that each was of "normal intelligence". All were adjudged to make effective use of hearing aids and of speech. A matched group of 24 normally hearing children served as controls. All subjects were from "average, middle-class schools and groups. None were rural or inner-city schools" (ibid., p.345).

The mean raw BTBC scores for the hearing-impaired children at 6, 7 and 8 years are 28.50, 32.33 and 35.25 whereas the comparable scores for the control group are 45.5, 48.33 and 48.58. The data were divided into two groups according to degree of hearing loss.

"While 33% of the milder hearing-impaired children scored above the 50th percentile, none of the severely impaired children scored at this level. Of the first group, 58% scored below the 10th percentile while 91% of the second group scored at or below the 10th percentile" (ibid., p.346).

Item analysis revealed erratic performance by the hearing impaired children (ibid.). Davis suggests this "may reflect early teaching emphasis or a lack of specific exposure to the concepts involved" (ibid., p.348). However, inspection of these children's responses (Appendix, ibid., p.351) casts doubt on Davis' assumption about "the sporadic, almost random nature of their responses (for example, only 55% of the children correctly answered 'next to' while 93% of them correctly answered 'nearest') (ibid., p.349). As will be shown later in the present investigation, Item 4 (next to) is harder than Item 16 (nearest) for many hearing children. Davis appears to have ignored the ceiling effect of a number of BTBC items for older middle-class children: one hundred per cent of her control group answered 33 of the 50 items correctly and only one item (pair) was answered correctly by less than 80%. 'Pair' also proved to be the most difficult for the hearing-impaired.

So judgements about erratic performance by the hearing-impaired do not seem to be warranted either by comparison with the control group or even by comparison with Boehm's (1971) norms for Grade 1, middle-socioeconomic groups. As Davis notes "those items yielding the fewest correct responses from the six-year-old children tended to yield a similar proportion of incorrect responses from the eight year olds" (ibid., p.348).

A tactile analogue of the BTBC has been developed. The Tactile Test of Basic Concepts (TTBC) was developed from Form A of the BTBC and a field test of its validity and reliability is reported by Caton (1977). The subjects, drawn from six residential schools and six day schools, were 75 blind children, 25 in each of grades K - 2, who either used braille or would eventually use braille. The 50 BTBC items were presented on plastic sheets as raised outline drawings which were simplified as much as possible. Whenever possible, simple geometric shapes were substituted for the BTBC pictures. All tests were performed tactually, even by those with residual vision.

Analysis of the performance of these blind children revealed a tendency for the easier items to be those which required the child to use himself as referent: items requiring comparative judgements were more difficult (ibid.).

"Concepts which were easier for blind subjects were: behind, next to, always, nearest, away from, and around. The indices of difficulty for these concepts ranged from .80 to .96. Concepts which were difficult ... were: half, middle, medium-sized, third, in order, pair, matches, and least. The indices of difficulty for these concepts ranged from .43 to .49. Concepts whose indices of difficulty fell between ... (sic) did not follow a specific pattern..." (ibid., p.384).

Unlike sighted children, the blind could not observe several figures simultaneously for comparison. The mean TTBC scores for each grade level were 32.00 (SD 6.52) for kindergarten, 35.16 (SD 7.00) for Grade 1 and 41.56 (SD 6.81) for Grade 2. A reliability (KR_{20}) coefficient of .87 was obtained indicating that the TTBC can be used with confidence with blind children (ibid.).

Comparisons of the performance of blind and sighted subjects were made. Differences in mean test performance as a function of grade level were evaluated by a 2 X 3 analysis of variance. Results "indicated that no significant difference was found between the two groups on the basis of visual status ($F = .672$, $df = 2$, $p < .01$) and a significant interaction was found ($F = 4.693$, $df = 2$, $p < .01$). Therefore it "was assumed that the tactile test did, in fact, perform the same function for the blind group as the print test for the seeing group" (ibid., p.385).

Caton suggests that the TTBC/BTBC concepts are ones which educators of the blind might tend to emphasise at an early age and hence the finding that the blind pupils knew the concepts almost as well as the seeing children. The blind kindergarten group performed slightly better than their sighted peers but the reverse was true for Grades 1 and 2.

Chi square analysis showed that, at kindergarten level the blind children, found the following items significantly <u>less</u> difficult than do sighted children: other, never, always, equal. 'Middle' was <u>more</u> difficult for the blind at kindergarten level whilst eleven concepts were more difficult in Grade 1 and twelve in Grade 2.

"The most striking result of the study is that no significant differences were found when the overall performance of blind children on the tactile test was compared with the overall performance of seeing children on the print test. Although it is not possible to say without further evidence that the tactile items are completely analogous to their print counterparts, this finding does substantiate the fact that the two forms of the test assess the same concept. Because of this, the Tactile Test of Basic Concepts is considered to be appropriate for use with children and is now being produced by the American Printing for the Blind. The major purpose of the test is to assist teachers in interpreting the test results of their students so that concepts which the students do not have can be taught at an early age" (ibid., p.386).

An exploratory investigation of the basic concept attainment of 45 educable mentally handicapped (EMH) children, aged 7:0 - 10:7 who had spent two or three years in school, employed the BTBC as the sole measuring instrument (Nelson and Cummings, 1981). The children's IQ scores ranged from 49 to 71 and all had passed vision and hearing screenings. Four of the BTBC concepts were failed by more than 75% of the children (in order, least, pair, third) and four more (equal, forward, left, medium-sized) by at least half the sample. An additional 17 concepts were failed by 25% - 49%. A developmental trend was established with the oldest group failing a mean of 10.21 concepts, the middle group a mean of 14.00 and the youngest group a mean of 20.33 ($F = 14.65$ $p < .001$).

"One question not answered by the present study is whether the EMH children totally lacked a conceptual understanding of the basic concepts or whether they merely did not comprehend the symbolic label" (ibid., p.306). The authors acknowledge the need for replication in various geographic regions.

CHAPTER 6

Review of Literature (iv): Urban and Rural Populations

Defining urban and rural populations is no easy task. "In a sociological context the terms rural and urban are more remarkable for their ability to confuse than for their power to illuminate" (Pahl, 1968). The word 'rural' has been used to describe

"those parts of a country which show unmistakable signs of being dominated by extensive uses of land, either at the present time or in the immediate past. It is important to emphasise that these extensive uses might have had a domination over an area which has now gone because this allows us to look at settlements which to the eye still appear to be rural but which, in practice, are merely an extension of the city resulting from the development of the commuter train and the private motor car" (G.B. Wibberley, cited by Clout, 1972 p.1).

However, from a sociological point of view

"it is possible to argue that rural people and rural communities may no longer be identified in developed parts of the world ... where all - or nearly all - inhabitants have now acquired urban-derived aspirations, enjoy urban life styles, and form parts of national systems of mass culture" (ibid., p.2).

The notion of a rural-urban continuum arose in reaction against the concept of two distinct ways of life but, warns Pahl,

"there are equal dangers in over-readily accepting a false continuity. Not only is there a whole series of continua but also there are sharp discontinuities, in particular the confrontation between the local and the national" (Pahl, op. cit., p.293).

Pahl asserts that "any attempt to tie particular patterns of social relationships to specific geographical milieux is a singularly fruitless exercise" (ibid., p.5). Owing to mobility and commuting "some people are in the country but not of it, others are of the country but not in it" (ibid., p.5).

Similarly, there is no clear division between inner city and suburb. Glass (1968) describes urbanism as "a baffling mixture of uniformity and diversity" and concludes that "we cannot unanimously trace its boundaries." Gans (1968) argues that no single urban, or suburban, way of life can be identified and that differences in ways of life between city and suburb are better explained by the class and life-cycle variations of their respective inhabitants.

It is a difficult task to show the extent to which "social areas are also coincident with local activity systems and possess a sense of place" (Herbert and Johnston, 1978 p.6). Many sociologists are critical of the concept of community neighbourhoods (Bell and Newby, 1978; Gittus, 1978; Gans, 1968). Herbert and Johnston argue that

"although the traditional form of the close-knit urban community has been modified over time it has not disappeared and that locality is a variable to be considered in the analysis of social interaction" (op. cit., p.7).

It is especially apposite to those with little mobility, in particular the very old, the very young and mothers with young children (ibid.).

"It is one thing to criticize the methodological inadequacies and shortcomings of 'community studies', but quite another to argue, or imply that we can do without studies of this kind" (Francis, 1983 p.120). Unfortunately we are largely dependent on community studies produced more than a decade ago which were generally lacking in conceptual clarity and sophistication (ibid.). A serious obstacle even now is that "we have not developed an adequate working conception of social and cultural organisation itself" (ibid., p.121).

"What is clear is that we need objective ways of getting at individual's and families' local social situations, their social milieux, in some better way than the overly contrasted blanket terms of 'isolation and anomie' versus 'community' (Bell and Newby, 1978 p.198).

Social geographers have assigned spatial boundaries to living conditions by considering the 'level of living' i.e. the actual conditions of life among the people in question or by their standard of living i.e. the aspiration levels of the people themselves (Knox, 1975). Knox considers the United Nations' Level of Living Index too blunt an instrument to measure variations in well-being in the UK. He formulated a working definition of the level of living for the British context:

"The level of living of persons resident within a given geographical area is constituted by the over-all composition of housing, health, education, social status, employment, affluence, leisure, social security and social stability aggregately exhibited in that area, together with those aspects of demographic structure, general physical environment, and democratic participation which may determine the extent to which the needs, the desires relating to the foregoing constituents of level of living can be, or are, met" (Knox, 1975 p.31).

Knox (1972) identifies fifty-three primary variables. Analysis revealed six components which are able to represent more than two-thirds of the variation. "Together the six components are taken as representative of the major dimensions of spatial variation in level of living in England and Wales..." (ibid., p.263). They demonstrate the influence of the processes of urbanisation and suburbanisation and the importance of the life-cycle.

"... it is clear that the most general similarities in the distribution of the primary variables involve a wide range of constituents of level of living which are in fact suggestive of urbanism rather than social class" (ibid., p.150).

Further analysis yielded four variables diagnostic of the major features of spatial variations in level of living: "the proportion of persons aged sixty years or more, the average number of persons per room, the proportion of private households in unshared dwellings without exclusive use of a fixed bath, and the proportion of economically active persons out of employment" (ibid., p.264).

By using these four variables Knox (1972, 1975) derives twelve types of local authority areas which illustrate spatial variations in level of living not only in degree but also in kind. For example, Type 2 is associated with a density of occupancy which is only slightly above average but with scores on the other three variables clearly below average. The twenty-six authorities comprising this type of area are mainly administrative counties and include Lindsay and the West Riding.[1] Type 3 authorities are associated with below average rates of occupational density and unemployment and above average proportions of persons aged 60 years or more. The East Riding is of this type. The authorities comprising Type 4 have an above average rate of unemployment and a below average proportion of poor

[1] All local authorities are those operating before 1974.

quality housing. Grimsby and Doncaster are included in this group. Apart from having a 'young' population, areas constituting Type 8 score poorly in relation to the major determinants of the spatial variations in level of living. Barnsley and Hull are of this type.

Index values, according to Knox, range from 9.7 (Berkshire) to 84 (Gateshead). The highest levels of living are in the Home Counties, Nottinghamshire (20.3) and Leicestershire (22.8) all of which are amongst Type 1. The lowest levels are recorded for Gateshead and South Shields (74.2), Type 10 and Wigan (78.9) which is grouped in Type 8. Barnsley and Hull, also Type 8, score 61.1 and 61.8 whilst Doncaster and Grimsby, both Type 4, have index values of 41.6 and 47.3. Lindsay and the West Riding (Type 2) score 42.1 and 49.7 whilst the East Riding (Type 3) has an index vlaue of 40.9. (Knox, 1972).

Though a useful instrument for mapping spatial variations in level of living, the index, especially in large urban areas, cannot portray the conditions experienced by the majority of people living there. The internal range of conditions is often as great as for the nation as a whole (Knox, 1972, 1975).

Recent social-indicator research casts doubt on whether there is any straightforward correspondence between material conditions, as measured by objective indicators, such as Knox used, and experienced life-satisfaction (Dale, 1980). Subjective indicators, based on direct reports from individuals about their own perception and feelings, may give information about life-domains that are more or less impossible to measure objectively. Unfortunately they may reflect biases arising from the method of interviewing and from the interpretation of answers. From an examination of case studies Dale concludes that

"it will be impossible to measure all aspects of the quality of life by a social-indicator approach. If the subjective evaluation and objective circumstances do not correspond, even when the criteria...are fulfilled then the explanation can be that there are important aspects of life which the indicators do not cover." (ibid., p.514)

"Every material improvement does not necessarily lead to an increased level of satisfaction. Nevertheless, the overall correspondence between the two sets of measures is probably greater than recent research might indicate. This seems especially to be true for the distribution of objective and subjective well-being within cities". (ibid., p.513)

Through a range of reminiscence, observation and reference Ward (1977), with great sensitivity, portrays various facets of life for the city child. He depicts the feel of growing up in the city highlighting those aspects of the city which are highly relevant to its children. For example, he cites a comparison of adults' and childrens' maps of the same area showing

"... a lighthouse which favoured as a significant landmark in all the maps drawn by adults. But none of the Harwich children showed the lighthouse on their maps, though many showed the public lavatory which stands at its base. Things which were important to them included kiosks, hoardings and other bits of unconsidered clutter in the street. One item which frequently recurred in their maps (and was totally ignored in those of the adults) was a telephone connection box - a large metal object on the footpath, with a fluted base. Obviously, as a feature for hiding behind or climbing on, this kind of obstruction has a value for children in their use of the street." (p.28).

Some of the places where city children play are identified:

"If you ask adults about their happiest or most vivid recollections of city childhood they will seldom talk about the park or the playground, but they will recall the vacant lot, the secret places behind billboards or hoardings" (p.87).

Thus, because "some bit of the city is designated as a play space on a plan, there is no guarantee that it will be used as such, nor that other areas will not be" (ibid., p.204).

Ward explores how some children show endless ingenuity in exploiting what the city offers while others remain isolated and predatory. He surveys the problems of privacy and space.

"The quest for personal privacy and the sense of social isolation are not opposites in the experience of the urban child. ... A survey conducted for the Community Relations Commission found that just under half of the children under five in the Handsworth district of Birmingham never went out to play. 'They have no access, either exclusive or shared, to play spaces at the front or back of the house and their parents feared for their safety if they let them out'." (ibid., p.42).

"The poor child, who is usually the most isolated from the life of the city as a city, is also, paradoxically, the child who is denied the solace of solitude" (p.47).

"Space, and the luxury of a room of one's own, are the positive advantages of the new or rebuilt environment" (p.40).

Perhaps the new environment of the council housing estate which is removed from the city centre is not always an improvement on the inner city as it is described by Clarke, Hall, Jefferson and Roberts (1976): The working-class neighbourhood

"which assumes its 'traditional' form in and after the 1880s, represents one, distinctive, example of the outcome of negotiation between the classes. In it, the different strata of the working-class have won space for their own forms of life. The values of this corporate culture are registered everywhere, in material and social forms, in the shapes and uses of things, in patterns of recreation and leisure, in the relations between people and the character of communal spaces. These spaces are both physical (the networks of streets, houses, corner shops, pubs and parks) and social (the networks of kin, friendship, work and neighbourly relationships). Over such spaces the class has come to exert those 'informal social controls' which redefine and reappropriate them for the groups which live in them: a web of rights and obligations, intimacies and distances, embodying in its real textures and structures 'the sense of solidarity, local loyalties and traditions'. These are the 'rights', not of ownership or force, but of territorial and cultural possessions, the customary occupation of the 'sitting tenant'. The institutions are, of course, cross-cut and penetrated by outside forces. The structures of work and workplace, near or far, link the local labour force to wider economic forces and movements. Not far away are the bustling commercial high streets, with their chain stores and super-markets, linking the home to the wider economy through trade and consumption. Through these structures, the neighbourhood is socially and economically <u>bounded</u>. At one level - the horizontal - are all those ties which bind spaces and institutions to locality, neighbourhood, local culture and tradition. At another level - the vertical - are those structures which tie them to dominant institutions and cultures. The local school is a classic instance of such 'double-binding'. It is the <u>local</u> school, next to houses, streets and shops where generations of working-class children have been 'schooled', and where the ties of friendship, peer-group and marriage are forged and unmade. Yet in terms of vertical relationships, the school has stood for kinds of learning, types of discipline and authority relations, affirmed experiences quite at variance with the local culture..." (Clarke et al., 1976 p.p.7/8)

The mixed blessings of urban renewal[1] are documented in a number of studies. Improved dwelling amenities, space and conditions are welcomed by tenants as are improvements in access to public transport but rents and

[1] Urban renewal is "an improvement in the physical environment of an area within an urban region, usually close to the centre, and a rearrangement of land-use patterns therein" (Jackson, 1980 p.225). Thus it entails more than just 'slum clearance'.

fuel bills increase (Jackson, 1980). The post-renewal areas are perceived as less friendly than the old ones (ibid.).

Misguided architectural designs of new 'urban villages' have, in many instances, separated the estates from the neighbourhood:

"... high density low-rise schemes in which, however cosy and intimate the imagery you are hardly ever aware of the presence of others. The Bartlett research has shown that walking about some experimental schemes you are as likely to meet another person in the middle of the day as you are in an ordinary residential street in the middle of the night" (Hanson, 1983).

The Bartlett School of Architecture research suggests that "the 'enclave' philosophy, however dressed up, will always create urban deserts" (ibid.).

Many new estates tend to obviate against a social mix of occupiers. Jackson and Marsden (1962) demonstrated that very many of the academically successful working-class children came from districts where social classes were mixed. Many housing schemes have also resulted in the segregation of the generations which is thought to have a destructive effect on the well-being of families and children (Bronfenbrenner, 1977; Rutter, 1981). Such schemes share a problem which is also common to large apartment blocks and high-rise flats where "large areas are neither private and therefore supervised by residents, nor truly public and therefore constantly used and overseen by passers-by." (Rutter, 1981 p.614; see also Hanson, op. cit.) Middle-class planners are accused of working on "the mistaken premise that you can create a home environment if you give people all the 'sanitary' necessities" (Pellow, 1981).

Furthermore, urban renewal policies have reduced the supply of low-rent housing both in the public and private sectors thus exacerbating the plight of the poorest who have to accept the worst dwellings (Field, 1977; Jackson, op. cit.; Murie, 1983). The least able of the poor are brought together by the local authority which places them in its least desirable estate (Evans, 1980). "As a result this estate will become even less desirable and tenants who are able to choose will move out of the estate... The result will be an estate marked by social problems of all kinds (ibid., p.205).

Meanwhile, delays in rehousing programmes intensify the appalling living conditions in scheduled slum clearance areas which suffer a rundown of services and cessation of repairs as they become 'doomed areas' (Gibson and Langstaff, 1982).

English, Madigan and Norman (1976) in their survey of slum clearance found many younger families who had never settled in the slum area knowing that their present accommodation was not going to be their final house. Though only a quarter to one third of middle-aged and older households were unequivocally against moving to alternative housing "the affectual ties of long acquaintance with the area and dependence on a network of established relationships were particularly important factors influencing their attitudes towards moving" (ibid., p.179).

"... the very limited benefits that have come from rehousing warn us both that the effects are likely to be marginal and also that the way rehousing is carried out may be crucial" (Rutter, 1981, p.618).

Relocation in estates peripheral to the inner city present their own problems. Many families end up "in cash terms, worse off than when they started once account (is) taken of the earnings of all members of the family, their travel costs, and the loss of earnings during periods of unemployment in the course of transition" (Jackson, 1980 , p.244). Male unemployment increases and there are fewer employment opportunities for women (ibid.).

Previous research on the effect of moving house has produced conflicting evidence (Hooper and Ineichen, 1979). A research project involving young families moving into a variety of new residential areas in Bristol finds a high rate of reporting of neurotic symptoms with wide variations between areas (ibid.). This study, it is claimed, confirms that "after an initial phase of difficulty, most families adapt to their new residential environment" (ibid., p.167). However, the picture is not so simple; neurotic symptoms were assessed within two years of the move and then again 18 months later: though the health of the wives who remained had improved there was no significant improvement of the remaining husbands. Moreover, families where the husband was the stressed member were more likely to have moved again than were families where the wife showed neurotic symptoms. The families who moved away were not followed up (ibid.). The highest levels of clinical neurosis, in both husbands and wives, were observed, at the first interview, in redeveloped inner-city houses and high-rise flats but not in central area maisonettes (Hooper and Ineichen, 1979; Ineichen and Hooper, 1974).

Comparisons of populations in an inner London borough and the Isle of Wight revealed higher rates of psychosocial problems and mental disorders in the city than in small town or rural communities (Rutter, 1981; Rutter and Quinton, 1977). "This is a real difference in the prevalence of disorder and not just an artifact of recognition or of differential migration" (Rutter, 1981 p.623).

"City living is associated with an increased susceptibility to a quite wide range of problems with different manifestations and different causes. ... cities are not all alike in having high rates of problems, nor are individual cities homogenous in this respect. ... the variations within cities are as great and as important as those between cities. These variations are not random but rather are systematically related to living conditions, in the design of housing complexes, and in the quality of schools (ibid., p.623).

Urban life for adults and children has been charted by the now classic studies of Young and Wilmot (1957) and Newson and Newson (1963, 1968, 1976). More recently, Wilson and Herbert (1978) explored the interrelationship between the parents' circumstances and the difficulties encountered by the children of fifty-six families known to the social services department of a large Midland city. Their findings include the following observation:

"The children's capacity to internalise events enables them to accept as normal a situation that, by any standard other than their own, would be considered extraordinary or tragic... We are not convinced that the lives of poor families in Britain constitute a 'culture' or a 'subculture', but we would argue that the variety of life-styles encountered among these families contains a tragic aspect. The draining of human energy and potential is the element that outweighs all others and has an overpowering effect on the growth of children that no other element can counterbalance" (p.104).

They refer to

"the feelings of hopelessness and powerlessness which all the families expressed at some time. Most of the families had a degree of awareness of their low status, and the low esteem in which they were held by others not in their group. There was an acute consciousness of the poverty trap and of the very limited opportunity structure in the market for jobs. Poor performance in the educational system was a parental memory reawakened by their children's experiences in school. Since upward mobility appeared obviously blocked, the only available solution to the pressures created by societal expectations was to develop a system of adaptive retrenchment. The needs of individuals must take second place. Decisions were made at family level and related to the main wage earner or recipient of benefit rather than to the needs of individual children. The value orientation of the child was thus shaped on a collective basis" (p.185/6).

This survival strategy of 'adaptive retrenchment' overrides attention to the individual needs of children.

"Parental attitudes are rooted in deep-seated feelings of powerlessness. Their view of the world is dominated by mistrust and there is no curiosity about it. Thus the processes of socialization reflect and reinforce parental feelings of failure and perpetuate inequalities within society." (p.186).

Ward (op. cit.) highlights one element in the situation of powerlessness: childhood reminiscence depicts the comparative ease of "flitting" in pre-war days. "At that time families were always moving. There were houses to let everywhere" (p.38). Ward points out that

"the element of freedom of choice in housing that their families had has totally disappeared. They would (now) be stuck in a particular bit of run-down accommodation in the fast-dwindling private sector, or if they were lucky they would be equally immobile in the flat the council had provided (as suitable for unsatisfactory tenants) waiting years for a transfer" (ibid., p.39).

The powerlessness of the urban poor in choice of housing is noted by English et al. (op. cit.) who are critical of many administrative practices and also by Ball and Newby (1978) and by Rutter (1981).

Relationships between school attainment and housing conditions have been established (Essen and Wedge, 1982; Essen, Fogelman and Head, 1978; Davie et al., 1972; Douglas, 1964).

"The implications of our findings on housing and attainment together with these studies of the environment are twofold: children who ever experience poor housing conditions, especially if this occurs throughout their childhood, clearly have lower attainment by the end of their compulsory school years. However it makes little difference how old they are when they experience these conditions. In addition, children will also have relatively low attainment if the area in which they live, as distinct from their own family, is at a disadvantage in terms of socio-economic characteristics" (Essen et al., 1977 p.57/58).

Children identified as disadvantaged at any one age are only a small proportion of all children who experience multiple disadvantages at some time in their childhood (Essen and Wedge, 1982). The National Child Development Study shows some delay in development on the part of adolescents who are no longer disadvantaged. Gan's explanation of deprivation, partly in terms of adaptations of behaviour and norms in response to stressful circumstances, together with the family's survival strategy of 'adaptive retrenchment' as described by Wilson and Herbert (op. cit.) is accepted by Essen and Wedge.

"This model of society appears particularly appropriate to our study as it provides a possible explanation of the continued delay in development of children who had been, but are no longer disadvantaged" (ibid., p.170).

Most of the literature on deprivation is concerned with urban areas. Relatively little focuses on rural deprivation. Rural deprivation is thought either not to exist or to warrant little attention or to be a local phenomenon (Cullingford and Openshaw, 1982).

"Because Britain does not possess areas equivalent to Oklahoma or West Virginia, the problems of the rural poor, or the unemployed, or the badly housed go largely unnoticed" (ibid., p.140).

"It is believed that the interpretation given to key indicatory variables is subtly different in rural areas, and thus social area techniques perfected in urban situations are considered to be inappropriate for rural applications" (ibid., p.411).

There is an urgent need for some means of identifying rural areas of deprivation. Cullingford and Openshaw suggest that the techniques of large scale surveys and local case studies should be combined to achieve this identification. This

"would also help to describe the differences between urban and rural deprivation, it would identify regional differences in the types of rural areas, and it would set particular rural areas in context both with respect to other rural areas and with urban areas" (ibid., p.416).

An Education Officer writes:

"In recent years many local education authorities have conducted objective surveys of deprivation within their areas in order to determine priorities for such programmes as the development of nursery education and the designation of schools as ones of 'special difficulty'. It has been interesting to note how the application of social indicators, of the sort pioneered in ILEA, have shown up rural communities as being as deprived as some of the larger conurbations. In one authority the indicators showed the most deprived area to be a small settlement, isolated, with a three teacher primary school and a high density of older council houses. The former rural district council had used the settlement for many families likely to default on rents or having damaged council property elsewhere, and the pattern of low expectation of behaviour had, through two generations, created a classically disadvantaged community. In another authority the indicators put a high priority on a small holiday coastal town in which there was an astonishingly high seasonal unemployment rate and a pattern of behaviour in which the breadwinner frequently left his family for the winter to seek employment..." (Centre for Information and Advice on Educational Disadvantage, 1980 p.10).

Since 1974, alterations in the distribution formula of the Rate Support Grant together with government aid under the Urban Programme have resulted in discrimination against rural and semi-rural areas (ibid.).

The same Education Officer avers that the "myth of the stimulating rural environment can easily be exploded in, for example, the Fenlands where land is so extensively farmed that there is less ground upon which the village child can play than his urban counterpart, ..." (ibid., p.106). This author identifies isolation as "the most significant disadvantaging factor for the rural child" (ibid.).

Sigsworth (1980) argues that urbanism has become the dominant culture and so "the development of formal education has been an essentially urban-based activity" (p.91).

"...rural features, while they have presented difficulties to the urban educational planner, wrestling with problems of inequality and disadvantage, have not until quite recently challenged his assumption of the essential correctness of applying urban scale to rural problems. Moreover, firm attachment to urban scale, when applied to a sparsley populated area immediately throws up the small rural school as a cultural abnormality requiring treatment" (ibid., p.91).

Citing the philosophy of the Swedish Lofoten Project, Sigsworth argues against an urban dominated curriculum which, "because it acts alienatively upon the pupil and destructively upon rural culture, disadvantages both" (ibid., p.96).

There is a vast amount of literature relating school attainment to social and ethnic disadvantage (for example, the literature arising from the NCDS) but virtually none which relates early school attainment to the type of area in which children live. In the years following the Plowden report, EPA schools were described (Midwinter, 1969, 1970; Halsey, 1972) and action-research programmes initiated and assessed (Halsey, 1972; Smith, 1975; Poulton and James, 1975; Woodhead, 1976). Since then the attention of research workers in Social Priority Areas seems to have focused on multi-racial problems. It would be interesting to find out how far and in what ways our inner city schools have changed since the early seventies.

With the exception of the West Riding EPA Halsey (1972) noted the relative inexperience, youth, high turnover and middle-class origins of EPA teachers. He also remarked on the "apparently lower interest of the EPA teachers in less able children". Low teacher expectations of disadvantaged pupils are well recorded (Pidgeon, 1970; Kelsall and Kelsall, 1971; Pilling and Pringle, 1978). Do the same attitudes and conditions appertain in SPA primary schools in 1983?

Given the current over-supply of teachers "the problem is likely to be less one of high turnover because of teacher's unwillingness to serve in 'difficult schools' and more one of the problem of maintaining staff morale and involvement in conditions of declining school rolls" (McKenzie, 1980, p.179). On the other hand, "falling numbers reduce stress and provide new opportunities for improving spatial standards" (Little, 1977 p.76). However, as Little points out, the concomitant is uncertainty about closure of schools.

The Plowden Report noted rapid population shifts in inner city areas resulting in ever changing school rolls. Urban renewal schemes may even have increased these fluctuations.

Little (op. cit.) observes a disillusion in some quarters "which argues that factors outside school are of such importance that what happens within schools can be of little effect" (op. cit., p.76). "There are wide variations in the quality of inner city education which any averaging tends to obscure" (ibid., p.77).

"Indisputably there are areas of outstanding and largely unrecognised success. ... Where there is failure the reasons are complex and mostly to be found within the school itself. Neither success nor failure can be deduced from the size or type of the school concerned" (ibid., p.79).

Rutter finds that several school features are strongly associated with children's problems viz.

"(a) a high proportion of children receiving free school meals (b) a high proportion of children of immigrant parents (c) a high teacher turnover (d) high pupil turnover (e) a high rate of absenteeism (f) a low pupil/teacher ratio" (Rutter and Quinton, 1977 p.182).

These characteristics were combined to form an index of school adversity

"The presence of school adversity produced no increase in psychiatric risk for children from disadvantaged families ... but it did considerably increase the risk with respect to children from non-disadvantaged families" (ibid., p.182; see also Essen et al., 1978).

Rutter (1981) identifies variations within secondary schools which considerably influence children's behaviour and attainments. Eight variables stand out as being of particular importance for successful schooling (ibid.).

"Firstly, it appears necessary that schools have a reasonable balance of intellectually able and less able children. Secondly, it has been found that the ample use of rewards, praise and appreciation tends to be associated with better pupil outcomes. Thirdly, a pleasant and comfortable school environment with well cared for and attractively decorated buildings is associated with better outcomes. Fourthly, ample opportunities for children to take responsibility and to participate in the running of their school lives seems conducive to good attainments and behaviour. Fifthly, children make better progress both behaviorally and academically in schools that place an appropriate emphasis on academic matters. Sixthly, pupils are influenced by the models of behavior provided by their teachers. Seventhly, it has been found that there are important skills involved in group management in the classroom. Finally, the evidence indicated that pupil outcomes tend to be better when there is a combination of firm leadership from the top and a decision-making process that involves all teachers and that leads to a cohesive approach in which staff members act together and support each other." (ibid., p.622).

Though the authors warn that it is premature, as yet, to draw firm conclusions, reports on the ongoing Strathclyde Experiment in Education point to promising outcomes from dynamic cooperation between parents and professionals in a severely deprived area (Wilkinson, Grant and Williamson, 1978). The Govan Project, situated in Moorpark Estate,

"is focussed on the socialisation systems used by adults in shaping children's lives. Adults in each component of this intricate network - the family, community and school - are often ignorant of the extent of their influence in the development of intellect and social awareness. By working with parents, community leaders, teachers and other professionals, we are attempting to bring about a deeper understanding of partnership in education" (ibid., p.2).

Wilkinson et al. comment on "the commonly held view that parents living in disadvantaged circumstances are either apathetic or antagonistic to their children's schooling. ... There is no doubt that apathy is widespread, but it would be wrong to conclude that parents do not care about the education of their children" (ibid., p.90). Responses to many of the project's events indicate a great deal of interest in education. "The difficulty would not seem to be apathy, but the lack of a structure through which this latent energy can emerge" (ibid., p.90).

"Educational leadership by local residents in inner city communities is a relatively new concept. Although the neighbourhood is familiar with other forms of leadership, (e.g. political and social) it has no experience of structural educational decision making. In several areas of the Project's work - preschool, family workshops and discussion groups, local leaders are beginning to emerge." (ibid., p.91).

Wilkinson et al. believe that

"education has a crucial role to play in the regeneration of inner city areas. It is through education that significant changes in attitude will be brought about - both in professionals and parents. ... We are convinced that no one formula for the eradication of deprivation exists. All we know is something of the nature of the ingredients to encourage local residents to throw off the cloak of apathy and to awaken in the professionals a new sense of service to the community that needs them" (ibid., p.3).

"If deprivation is to be counteracted, it is essential that the potential of both professional and parent be coupled with adequate resources and political support" (ibid., p.92).

CHAPTER 7

The Present Investigation

In the light of the evidence and theories reviewed in Chapters 3, 4, 5 and 6, together with the arguments presented in Chapter 2, objectives and working hypotheses have been formulated.

The objectives are

1 To provide standard BTBC (Form A) scores for each age group, of children in England, by half years between 3½ and 7½ years.

2 To produce a shortened version of the BTBC suitable for children from 3½ to 5 years.

3 To ascertain the need for provision of two sets of standard scores for each age group, one for use during the first half of a school year and one for use during the second half, by testing the null hypothesis:

H_0 There is no significant difference between the mean scaled BTBC scores (Form A) of children tested during the first half of their school year and those tested during the second half (as defined by September to the start of the February half-term break and the end of the February break to the end of the summer term for all children on roll in September. For children admitted to nursery or reception classes in January the second half of their school year is defined as starting on June 1 and for April admissions on November 1.)

4 To provide evidence of the reliability of the BTBC (Form A).

5 To test the equivalence of Forms A and B of the BTBC.

6 To provide evidence of content validity by

a examining how frequently the basic concept terms of the BTBC appear in teachers' classroom discourse,

b word counts of commonly used early reading schemes,

c word counts of commonly used first school mathematical texts,

d word counts of "Science 5/13" texts,

e perusal of 'handbooks' for nursery teachers.

7 To provide some evidence of predictive validity by comparing BTBC scores with teachers' assessments of pupils' achievements in spoken language, reading and mathematics one year after they have been tested with the BTBC (Form A). The unreliability of such criteria is acknowledged and will be discussed later.

8 To provide evidence of the construct validity of the BTBC by principal components analysis.

9 To ascertain any sex differences in the BTBC results by testing the following null hypothesis:

H_0 There are no significant differences between boys' and girls' mean scaled scores on the BTBC.

10 To test the following hypotheses concerning social class differences:

H_0 There are no significant differences between the BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations), group 2 (as defined by parents in other 'white collar' and skilled manual occupations) and group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment).
(See Review, pp. 75-82, 110; Results p. 266)

H_1 The BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations) are significantly higher than those of children in socio-economic group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment). (See Results, pp. 266-7)

H_2 The BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations) are significantly higher than those of children in socio-economic group 2 (as defined by parents in other 'white collar' and skilled manual occupations). (See Results, pp. 266-7)

H_3 The BTBC mean scaled scores of children in socio-economic group 2 (as defined by parents in other 'white collar' and skilled manual occupations) are significantly higher than those of children in socio-economic group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment).
(See Results, pp. 266-7)

H_4 Social class differences (as defined by parental occupations), which are reflected in the BTBC mean scaled scores, are greater below 5 years than at 7 or 7½ years of age.
(See Results, pp. 267-9)

11 To test the following hypotheses concerning various kinds of domiciliary areas:

H_0 There are no significant differences between the BTBC mean scaled scores of children attending rural, inner city or 'other urban' schools (as defined by the DES, 1978).

(See Review, pp. 138-152; Results, p. 270)

H_5 The BTBC mean scaled scores of children in socio-economic groups 2 and 3 (as defined by parental occupations) attending rural or urban schools in areas with marked social difficulties (as defined by low incomes, high unemployment and lack of amenities) are lower than those of children in socio-economic groups 2 and 3 attending schools in areas without a preponderance of such difficulties. (See Results, pp. 270-72)

After observation in schools the following hypothesis was also proposed:

H_6 The BTBC mean scaled scores of children attending inner city schools (as defined by the DES, 1978) will be higher than those of children attending schools in urban rehousing estates (as defined by post-war slum clearance).

(See Results pp. 272-3)
(See also p. 274)

12 To test the following hypotheses concerning children's <u>non-linguistic strategies and their understanding of relational concepts</u>:

H_0 There is no significant difference between the number of correct responses to a BTBC term which is unmarked (positive) and the number of correct responses to the opposite marked (negative) term, as defined by Clark (1977).

(See Review, pp. 62-72; Results, pp. 275-6)

H_7 The number of correct responses to a BTBC term which is unmarked (positive) is significantly higher than the number of correct responses to the opposite term which is marked (negative), as defined by Clark (1977).

(See Results, pp. 275-6)

H_0 There is no significant difference between the frequency of (a) the larger object (or greater amount) chosen in error and (b) other errors on each BTBC item where this type of error is possible.

(See Results, pp. 277-9)

H_8 The frequency of (a) the larger object (or greater amount) chosen in error is higher than (b) other errors on each BTBC item where this type of error is possible.

(See Results, pp. 277-80)

H_0 On each item 34 and 41 (Form A) there is no significant difference between the wrong responses chosen.

H_9 On each item 34 and 41 (Form A) the frequency of wrong responses on the vertical axis is higher than the frequency of wrong responses chosen from the horizontal axis.

(See Review, p. 64; Results, pp. 280-81)

H_0 On item 29 there is no significant difference between the wrong responses chosen.

(See Review, p. 64; Results, p. 281)

H_{10} On items 26 and 29 the frequency of the topmost object as a wrong response is higher than any other error frequency.

(See Review, pp. 62-72; see Results, p. 282)

H_0 On item 9 there is no significant difference between the wrong responses chosen. (See Results, p. 283)

H_{11} On item 9 the frequency of the nearest object as a wrong response is higher than any other error frequency.

(See Review, pp. 62-72; Results, p. 283)

H_0 There is no significant difference between the number of times the first object or picture in a row is chosen as an alternative wrong response to the BTBC items than the number of times another alternative wrong response is chosen.

(See Review pp. 91-93; Results, pp. 283-86)

H_{12} The number of times the first object or picture in a row is chosen as an alternative wrong response to a BTBC item is significantly higher than the number of times any other response is chosen.

(See Results, pp. 283-6)

CHAPTER 8

The Method

i The Sample Population

ii The Boehm Test of Basic Concepts (BTBC): Technical Data

iii The Procedures

iv The Statistical Analyses

CHAPTER 8

The Method. i. The Sample Population

1.00 **Location**

Following advice from the NFER, through correspondence and conversation, the population sample was drawn from one non-Metropolitan and two Metropolitan areas. Humberside, Barnsley and Doncaster were conveniently accessible.

1.01 These areas represent four of the twelve types of levels of living identified by Knox (1972, 1975). Their level of living index values are in the middle ranges (ibid.). Table 1 shows their positions relative to the extremes of the index.

	Type	Level of living index value
Buckinghamshire	Type 1	9.7
- - - - - - - - - -		
East Riding (partly absorbed by Humberside)	Type 3	40.9
Doncaster (pre-1974 CB)	Type 4	41.6
Lindsey (partly absorbed by Humberside)	Type 2	42.1
Grimsby (pre-1974 CB; now in Humberside)	Type 4	47.3
West Riding (partly absorbed by Barnsley, Doncaster and Humberside)	Type 2	49.7
Barnsley (pre-1974 CB)	Type 8	61.1
Kingston-upon-Hull (pre-1974)	Type 8	61.8
- - - - - - - - - -		
Gateshead	Type 10	83.9

Table 1 Types of level of living pertaining to the sample.

1.02 It is recognised that Knox's figures are now outdated but it is unlikely that the extent of social disparities between the twelve types of area has reduced. Whilst the level of living index and type classification is a useful guide it must be remembered that the internal range of conditions in any one area is often as great as for the nation as a whole (ibid.).

1.03 Expense and time precluded wider sampling.

1.04 Since 1974 the three areas have inherited buildings, equipment, administrative, teaching and ancillary staff from several LEAs. Thus, they exhibit a range of ethos, teaching styles and facilities.

2.00 The Schools

The distribution, in 1980, between the three LEAs, of schools containing children whose ages fall between three and eight years, together with the total population figures for each area (1979 mid-year estimates) is as follows:

	Nursery only	Infant	First	I/J*	F/M*	N	Total population
Barnsley	5	28	7	57	0	97	221,800
Doncaster	0	22	49	23	13	107	286,500
Humberside	10	45	112	150	4	321	849,250
	15	95	168	230	17	525	

Table 2 Distribution of schools between LEAs.

* I/J: Infant-Junior; F/M: First-Middle

2.01 With a sampling factor of 1/7 and an administrative factor of 1/6 so as to allow for refusals etc., 88 schools were drawn at random by numbering the schools down the LEAs' schools lists and using a random numbers table (Haber and Runyon, 1979 p.A59).

2.02 Of these 88 schools, (1/6 of the total), two closed before they could be visited and two could not cooperate; one because of staffing difficulties and one because of major building reconstruction. Twelve were unvisited as their distance from base rendered the journeys involved uneconomical of time and petrol.

2.03 Six further schools were added to the remaining 71. These were inner city schools which were added to make the sample more representative of the nation's schools. They were selected by numbering the inner city schools not already included in the sample and drawing the numbers from a box.

2.04 The final sample of schools which were used is approximately 1/7 of the total number of schools and is distributed as follows:

	Nursery only	Infant	First	I/J	F/Middle	N
Barnsley	0	7	1	9	0	17
Doncaster	0	0	17	1	2	20
Humberside	4	6	20	10	1	41
	4	13	38	20	3	78

Table 3 Distribution of schools sampled.

2.05 The sampling achieved an acceptable balance of schools between the three LEAs.

	No. of schools in 3 - 8 age range	No. of schools in sample
Barnsley	97	17
Doncaster	107	20
Humberside	321	41
	525	78

Table 4 Distribution of sampled schools between LEAs.

The differences between the expected and actual distributions are non-significant (Chi square 2.456 df 2).

2.06 Rural and urban areas are difficult to define (see Chapter 6). The arbitrary definitions and the methods of classification used by HMI were followed:

"The sample was not stratified by locality because no definitions of particular schools by locality could reliably be made prior to HMI visits. HM Inspectors defined schools by locality at the time of the inspection; ...

Inner city: the centre of large conurbations and the inner rings of large cities.

Rural: hamlets, villages and small towns with a population of 15,000 or less.

'Other urban': any area not able to be categorised as inner city

> or rural, including towns exceeding 15,000 and certain parts of cities and conurbations." (DES 1978 p.155)

Further elucidation was provided by correspondence with the DES. A member of the Inspectorate, who took part in the Primary Survey, explained how she would have classified certain local schools supposing they had been used in the survey. Thus models were provided.

2.07 The distribution of schools in the sample across rural, inner city and 'other urban' areas parallels the distribution found by HMI (Chi square .139 NS d.f. 2).

	No. of Schools					
	Barnsley	Doncaster	Humberside	N	%	HMI Primary Survey
Rural	14	8	9	31	40%	39%
Urban	3	12	20	35	45%	45%
Inner City	0	0	12	12	15%	17%
	17	20	41	78		

Table 5 Distribution of rural, urban and inner city schools.

2.08 A range of school sizes was distributed as follows:

No. on roll*	< 50	50 - 99	100 - 199	200 - 299	> 299	N
No. of schools	4	11	37	19	7	78

* equivalent full-time places in the case of nurseries.

Table 6 School sizes.

The smallest school was a rural 5 - 11 school with 16 pupils and the largest was a rural 3 - 8 school with 330 on roll.

2.09 Sampling within schools

Stratification by socio-economic grouping was impossible to do in advance. Schools do not always have this information and are generally reluctant to seek it from parents. It was expected that a selection of children by birth dates would produce a representative cross-section of socio-economic groupings and ability levels. Evidence from the British Ability Scales Sampling population confirms this expectation (Elliott, 1983).

2.10 The following categories of children were excluded from the sample:

a Children who live in homes where English is rarely spoken.

b Deaf or partially-hearing children.

c Children with visual difficulties not compensated for by spectacles.

d Children known to have suffered brain damage.

e Children attending special schools or awaiting entry to a special school.

f Children diagnosed as ESN(S) or ESN(M) or borderline cases.

2.11 Two categories of children which ideally should have been included in the sample, but were not, were three and four year olds not attending nursery schools or classes and children in private education. Home visiting of preschool children would have been too time-consuming. There are very few private schools for the under-eights in the areas used. For instance, there are only four small private infant schools in Doncaster Metropolitan District.

2.12 It was anticipated that three boys and three girls whose birth dates fell closest to and within a month of the day of testing, making 6 children of each age 4, 5, 6 and 7 years, would be selected from each school visited. Similarly three boys and three girls whose birth dates plus six months fell closest to the day of testing making their ages 3:6, 4:6, 5:6, 6:6, and 7:6 would be selected. Falling rolls often made it impossible to achieve an exact balance of boys and girls or six of each age group in every school. The rule regarding birth dates was never violated.

2.13 A total of 1928 children were tested with Form A: 965 boys and 963 girls. Analysis of variance confirmed equivalent distribution of boys and girls between age groups. (F = .299 df 9/1918)

2.14 According to whichever parent's occupation ranked highest each child was classified by socio-economic group as follows:

Group 1: professional (including nurses and teachers) and managerial (i.e. the Registrar-General's classes I & II).

Group 2: other 'white collar' and skilled manual (i.e. the Registrar-General's class III).

Group 3: semi-skilled and unskilled manual (i.e. the Registrar-General's classes IV & V) and parents in long-term unemployment.

Group 4: known to be manual workers but the precise nature of the job unknown i.e. groups 2 or 3.

Group 5: known to be non-manual but insufficient information to distinguish between groups 1 and 2.

Group 6: occupation unknown.

This coarse grading enabled classification where the information was not sufficiently precise to use the Registrar-General's classification; for example, "labourer" could be placed in Group 3 even though it was not known whether the man was a bricklayer's labourer (IV) or a navvy (V). Some errors may have been made when classifying miners: all were graded as underground workers unless it was known that they worked above ground. It is argued that the tight-knit mining communities produce the same kind of homes for their children regardless of the precise nature of the father's job. The OPCS classification of occupations was otherwise adhered to.

2.15 The final sample was distributed across socio-economic groups as follows

Groups	1	2	3	4	5	6	Total
N	325	721	452	139	19	272	1928
Sample percentage	16.9	37.4	23.4	7.2	1.0	14.2	

Table 7 Distribution of socio-economic groups.

2.16 Analysis of variance confirmed homogeneity of age distribution across the six groups ($F = 1.16$; df 6/1920 NS).

2.17 Analysis of variance confirmed equivalent distribution of boys and girls in each socio-economic group ($F = 0.54$; df 6/1921 NS).

2.18 Table 8 shows population estimates compared with the distribution of socio-economic groups where there is sufficient information available about the parents' occupations to place subjects in Groups 1, 2 or 3. The population estimates are based on the number of persons in employment in 1977 (CSO, 1980). The sample figures also include the long-term unemployed; therefore comparison cannot be made with any accuracy. Nevertheless, the sample distribution shows a similar trend to the national distribution of socio-economic groups.

Registrar-General's classes:	I & II		III		IV & V	
	Males	Females	M	F	M	F
Population estimate:	23	6	56	61	21	33
Sample percentage (N = 1498):	22		48		30	

Table 8 Population estimates and sample distribution by SES.

2.19 A sample of the range of parental occupations can be seen in Appendix 8.

The details given in this subsection are all taken from the BTBC Manual (Boehm, 1971). A copy is in Appendix 4.

1.00 Forms A and B of the BTBC were designed as group tests, with the option of individual administration, for children in American kindergartens and grades 1 and 2.

1.01 The purpose of the BTBC is "to identify

(a) individual children whose overall level of concept mastery is low and who therefore may need special attention, and

(b) individual concepts with which large numbers of children in a class may be unfamiliar.

In this way the test is intended both as a "detector" and as an instructional device for use by the classroom teacher" (Boehm, p.4).

1.02 Boehm classifies 23 of the 50 BTBC terms (in the form in which they are tested) as space concepts, 18 as concepts of quantity, 4 as time concepts and 5 as miscellaneous. Two terms obviously include more than one concept: 'some, not many' (item 6) and 'not first or last' (item 32). (The classification and multiplicity of concepts in these and other items will be discussed in later chapters.) The 50 BTBC concept terms are listed in Appendix 1.

1.03 The BTBC consists of two booklets for each Form. Copies of the Form A booklets are in Appendix 2 and Form B booklets in Appendix 3. Each booklet contains 25 sets of pictures with 3 practise sets on the front cover. The subject has to select the picture indicated by the instructions. For example, item 3 of Form A shows a table and three boxes, one box is on the table, one under it and the third is approximately 6cm. away from it. The direction is: "Look at the table and the boxes. Mark the box that is away from the table."

Detailed directions are given for each item. The directions for both Forms are in the Manual in Appendix 4.

1.04 Class record forms are available. A copy is in Appendix 5.

2.00 Standardisation of Form A

"The BTBC was designed as a screening and teaching instrument rather than for predictive or administrative purposes. Consequently, it was considered unnecessary to select standardization samples representative of children in kindergarten and the first and second grades in the nation as a whole." (Boehm, 1971, p.19)

Children in the Form A standardisation sample, which was the basis for "beginning-of-year" norms, were from schools in 16 cities located across the United States whilst the children in the sample for "midyear" norms were from schools in five cities.

2.01 "School officials in each cooperating city were asked to provide classroom groups from schools with a fairly wide range of socioeconomic background. Selection was to be based on the socioeconomic level of the primary areas from which the schools drew their enrollments. No formal specifications for this selection were given; the choice was left to the judgment of the administrative personnel." (Boehm, 1971 p.19)

Inspection of Table 3 of the Manual (standardisation samples by city, grade and socio-economic level) reveals discrepancies between socio-economic levels for "beginning of year" testing and "midyear" testing. These discrepancies are shown below.

Kindergarten	Beginning of year	Midyear	Beginning of year	Midyear
socioeconomic level	N	N		
Low	1921	162	54.6%	18.7%
Middle	912	453	25.9%	52.4%
High	684	250	19.4%	28.9%
Grade 1				
Low	2303	276	49.4%	27.9%
Middle	1313	435	28.2%	43.9%
High	1043	280	22.4%	28.3%
Grade 2				
Low	824	222	52.8%	27.3%
Middle	381	349	24.4%	42.9%
High	356	242	22.8%	29.8%

Table 9 BTBC standardisation samples by school grade and socioeconomic level.

Table 9, which is based on Table 3 in the Manual, shows that higher proportions of children at a low socio-economic level were used for the "beginning of year" testing than for "midyear" testing.

3.00 Alternate Forms

Two alternate forms of the BTBC are available. "Form B was designed to measure knowledge of the same concept as that measured by the corresponding item of Form A" by attempting "to present each item in a context similar to that of the Form A item." (Boehm, 1971 p.19)

"An 'armchair' attempt was also made to match the difficulty of Form B items to that of the corresponding items in Form A. Thus the two forms of the BTBC were designed to parallel each other in both coverage and difficulty level." (Boehm, 1971 p.19)

3.01 Equivalence of Forms A and B

A comparability study of Forms A and B found an overall comparability of scores: the mean scores were 42.4 on Form A and 42.9 on Form B. (N = 1192 SD = 7.3 Form A, 7.0 Form B)

3.02 "Frequency distributions of the raw scores were prepared separately for the two forms, and percentile equivalents were computed for each raw score. A graph was prepared for each form, showing the relationship between each raw score and its corresponding percentile. Perfect equivalence of the forms would have been indicated had the two graphs overlapped throughout their range. While the graphs did not correspond perfectly, at no point were scores on Form B more than 1.5 points higher than the comparable scores on Form A." (Boehm, 1971 p.21)

These graphs are not presented in the Manual. Tables showing the percentage of children passing each item by school grade and socio-economic level are provided for each form. Those for Form A are based on the Form A standardisation sample. Those for Form

B are "estimated from corresponding data for Form A based on results of equivalence study of Forms A and B." (Boehm, 1971 p.24/25)

4.00 Norms

The Manual presents two tables for each form: one showing the percentage passing each item by grade and socio-economic level at the beginning of the school year, the other at midyear. Data for beginning of year norms was obtained in September and October whilst midyear data was obtained between mid-November and late February.

The "midyear data reflect a generally greater mastery of concepts than is found at the beginning of the year... The exceptions occur mainly at the middle and high socioeconomic levels of grade 2, where the items tend to be quite easy regardless of the month of testing." (Boehm, 1971 p.21)

4.01 Tables 9 and 10 of the Manual present percentile equivalents of raw scores, by grade and socio-economic level, derived from the Form A standardisation sample. The mean raw scores shown below are extracted from Table 9 which relates to beginning of year testing.

	Kindergarten			Grade 1			Grade 2		
	Socioeconomic level								
	Low	Middle	High	Low	Middle	High	Low	Middle	High
Mean	25.5	31.8	35.8	33.8	42.0	43.7	41.2	46.9	47.3
N	1921	912	684	2303	1313	1043	824	381	356
SD	8.9	8.6	7.9	8.9	5.4	4.9	6.3	2.9	2.7

Table 10 BTBC mean raw scores for beginning of year testing.

The mean raw scores shown below are extracted from Table 10, in the manual, which relates to mid-year testing.

	Kindergarten			Grade 1			Grade 2		
	Socioeconomic level								
	Low	Middle	High	Low	Middle	High	Low	Middle	High
Mean	28.4	35.3	39.4	39.2	43.8	45.6	43.5	46.7	47.8
N	162	453	250	276	413	280	222	349	242
SD	8.1	8.0	6.5	5.5	4.5	3.7	5.0	2.7	2.6

Table 11 BTBC mean raw scores for mid-year testing.

4.02 "Since total scores on Forms A and B were found to be equivalent, the norms presented in Tables 9 and 10 may be employed for total scores obtained on either Form A or Form B." (Boehm, 1971 p.26)

5.00 Reliability

Table 11 of the Manual presents split-half reliability coefficients and standard errors of measurements for each Form by grade and socio-economic level. Form A reliability sample consists of the 2647 children whose scores were used in deriving the mid-year norms. Form B reliability sample consists of approximately one thousand children in the equivalence study who were administered Form B before Form A.

5.01 <u>The split-half reliability</u> coefficients, corrected by the Spearman-Brown formula, for the total score on Form A range from .68 to .90 while the corresponding coefficients for Form B range from .12 to .94. "The coefficient of .12 is, of course, quite low. It was obtained for the grade 2, high-socioeconomic-level sample, a group which had a mean total score of 48.5 and a standard deviation of 0.9. Such a group forms the extreme upper end of the ability groups for which the BTBC has been presumed applicable. At this level, the value of the BTBC would seem to lie only in the identification of children who are far below the group's average ability." (Boehm, 1971 p.28)

5.02 <u>The standard errors of measurement</u> range from 1.4 (grade 2, high socio-economic level) to 3.0 (kindergarten, low socio-economic level) for Form A and 0.9 (grade 2, high socio-economic level) to 3.4 (kindergarten, low socio-economic level) for Form B. For the total groups of pupils the SE_m for Form A was 2.7 for kindergarten, 2.1 for grade 1 and 1.7 for grade 2. For Form B the SE_m was 3.0 for kindergarten, 2.2 for grade 1 and 1.7 for grade 2.

"The standard errors of measurement obtained for BTBC total scores appear to be essentially comparable for Form A and Form B." (Boehm, 1971 p.28)

5.03 <u>Alternate form reliability</u> is reported for total scores on Forms A and B by socio-economic level and school grade. The coefficients are based on all children in the equivalence study who were tested on both forms. The coefficients range from .58 (kindergarten, low socio-economic level) to .92 (grade 2, low socio-economic level). The coefficient for the total group of pupils at kindergarten level is .72 (N = 291), at grade 1 it is .87 (N = 423) and at grade 2 level .88 (N = 421). The latter group includes 151 children who were inadvertently tested with Form A first instead of Form B.

6.00 Validity

No evidence of validity is given in the Manual. (Validity studies are reviewed in Chapter 5.)

1.00 The Standardisation of Form A

Each child was tested individually for two main reasons. Firstly, although group testing might appear to be infinitely quicker, it is not always so in practise with young children: the time taken to organise tables facing one way, to check that all the children are looking at the right row of pictures, to allow for someone needing a toilet visit, to work at the pace of the slowest can occupy a large amount of time to say nothing of the disruption of the class and school day and the difficulty of finding an unoccupied room. Secondly, group testing can be unreliable with young children: they are used to working cooperatively so that copying could be rife. Moreover, at the lower end of the ability scale or with the youngest children a group test may merely distinguish between those who can follow the general instructions and those who cannot. Group testing also requires more than one supervisor.

1.01 An English adaptation of the test instructions was effected as follows:

Item 4: 3 and 4 year olds were also given the alternative word 'lorry' as well as 'truck'

Item 8: 'cakes' replaced 'cupcakes'

Item 22: 'bricks' replaced 'blocks'

Item 30: 'puddings' replaced 'desserts'

Item 33: 'biscuits' replaced 'cookies'

Item 40: 'sweets' replaced 'candies' (See Appendices 2, 4 and 6)

The first 300 children tested were each given the original BTBC form of the items listed below, then, if they failed, they were offered an alternative version. This was done without the children realising their first answer had been recorded or was incorrect.

Item 15: '... the cake that is whole' is not a common English usage. Proportions passing this item were 34% of 3-4 year olds, 44% of 5 year olds and 61% and 73% of 6 and 7 year olds. When 'the whole cake' was substituted the proportions passing changed to 59% of 3-4 year olds and 75%, 92% and 98% of 5, 6 and 7 year olds respectively. So both forms were used in the directions for item 15: "Show me the cake that is whole, the whole cake."

Item 40: British children encounter 'Zero' as a prelude to blast-off but rarely as an adjective. The proportions of children responding correctly to "... the box with zero sweets" were 5% of 3-4 year olds, 11% of 5 year olds and 30% and 33% of 6 and 7 year olds. The proportions responding correctly to "... the box without any sweets" were 74%, 97%, 97% and 100% respectively. So the words 'without any' replaced 'zero' for subject 301 and all subsequent subjects.

Item 46: 'skip' is seldom used in this context with very young English children. No 3, 4 or 5 year olds, in the sample of 300, responded correctly to 'skip' and only 13% of 6 year olds and 35% of 7 year olds did so. When 'miss' was substituted 49% of 3-4 year olds passed this item and 55%, 85% and 91% of 5, 6 and 7 year olds passed. So 'miss' replaced 'skip' for subject 301 and all subsequent subjects.

Item 37: It was observed that only the 7 year olds responded adequately to 'medium-sized'. 'Middle' is a common term in nursery and infant classes but 'medium' is rarely used except in mathematics with the older first school children. So 'middle-sized' replaced 'medium-sized' for subject 301 and all subsequent subjects younger than seven. Both terms were used with 7 year olds: "... the fish that is middle-sized, the medium-sized fish."

Items 27 and 47: a bright, confident seven year old who scored 50 suggested slight alterations to the wording of these two items. When asked, after the test, why he had hesitated over these items he suggested that the words should have been "... as many marbles as the box has" for item 27 and "... two pictures that have equal numbers..." for item 47. So for subject 104 and all subsequent subjects his suggestions were incorporated.

Item 23: the wording of this item was simplified. (See Appendix 6.)

Slight, non-substantive alterations were made to the wording of items 1, 5, 9, 10, 12, 13, 18, 19, 21, 26, 27, 28, 38, 41, 42, 45, 50.

The altered directions for Form A are in Appendix 6 and the BTBC (1971) directions are in Appendix 4.

1.02 Adaptations for individual testing were made in an attempt to make the situation more natural and informal without altering the essential nature of the standard BTBC directions. Instead of prefacing the instruction for every item with "Look at the..." the interrogative phrase "See the...?" was sometimes substituted.

Children were asked to "Point to..." or "Show me..." or "Find..." or "Touch the...". In cases where pointing became well established during a test such words were then omitted. This happened with most of the school age children.

For item 34 the subject was asked to "Put your finger below the table" and for item 46 "Put your finger on the box with the X, this one. Now miss a box and point to the next box. Miss a box and point to the next box."

All these adaptations are written into the altered directions in Appendix 6.

1.03 From previous experience with the BTBC (Smith, 1977) the practise items were deemed unnecessary in individual testing. Very few children failed item 1. For those that did so the first item was repeated after item 50. Only one child changed her response.

1.04 School Visits

A preliminary visit was made to eleven of the schools in the drawn sample. (See Section i, 2.01 - 2.04) The others were first contacted by telephone. Headteachers were told that research into young children's language development was in progress and were asked if the investigator might visit and discuss their school's possible participation. The word "test" was avoided at this stage unless the headteacher sought further information whereupon details were given. Reactions ranged from deferral of a visit for various cogent reasons to "You are welcome to come at any time."

One infant headteacher replied: "Today? I've got 32 NNEB students visiting and a puppet show for the whole school and two teachers absent. But why not? One more won't matter. Come about 11.15. The puppeteer should have finished by then."

At the first visit the headteacher and any interested staff were given details of the research project and what was requested from the school i.e. access to class lists or registers and a small space in which to test without distractions for the child (cloakroom, hall, office or empty classroom etc.). The wariness detected in a few headteachers' voices over the telephone vanished during these conversations. All evinced a warm welcome, complete cooperation and generous hospitality.

The sample of children from that school was then drawn from the class lists or registers. (See Section i, 2.09 - 2.12) Arrangements were made to start testing immediately or, in some cases, to return on a day more convenient to the school.

1.05 The order and timing of visits to the 78 schools which were used was determined by economy of travelling time and by occasional school holidays, trips and other school functions, Christmas and end-of-term activities, school decorators and an epidemic of mumps.

Only in very small schools was it possible to visit the headteacher and complete the testing there in one day. Several days were generally required and at least a full week for the largest schools.

Where there were few children in the school whose birth dates rendered them eligible for the sample, a further visit was made some weeks later.

1.06 <u>Testing the Children</u>

The teachers were asked not to mention the word 'test' in the children's hearing. Each school child was asked to help with some puzzles or games which were being tried out "to see which ages of children they were best for". The nursery children were simply invited to "come and play a game with me."

With very rare exceptions the children were all eager. The handful that were reluctant were told "Alright, if you change your mind you can come later.": they did. Only two out of 1931 children steadfastly refused; one a nursery child, the other a five year old. Testing was abandoned with another nursery child who was emotionally disturbed and never spoke in the nursery. Though he showed some interest in the BTBC pictures he was unable to concentrate.

With the school children it was explained that certain children had been chosen because of their birthdays and apologies were made to those left out. Any very disappointed children in the reception and nursery classes were given 'turns'. i.e. a shortened version but not included in the standardisation sample.

1.07 The children were tested in whatever order was most convenient to the school. Care was taken not to interrupt stories or assemblies. No child was taken away from play in which he was totally absorbed and which would have been destroyed by interruption.

1.08 Rapport was easily established and the children enjoyed the test. Each child was seated on the examiner's left so that recording could be surreptitious. No indication was given of right or wrong answers. Encouragement was maintained throughout by such remarks as 'Great', 'Good', 'Yes', 'Smashing', 'You are good at this game', 'Well done' or, with surprise, "I didn't think girls of five would be able to do that one." No comment was made on obvious or suspected guesses. Wrong responses, where the child seemed to know it might have been wrong, were greeted neutrally with 'Um', 'O.K.' or silence. When a child showed complete puzzlement or said "don't know" the examiner gave reassurance by saying "That's alright. Hardly anyone gets them all right" or "That one's really for much older boys" or "Let's leave that one. It's too difficult. Do this one." Some of the youngest children showed their awareness of difficulty by trying diversionary tactics such as "Do you know my Daddy? He's going to...".

Every child was thanked and left the examiner feeling she had done well. As a rule the only children who knew their actual score were the thirty subjects who scored 50. They were congratulated on having got every one right. Several older children asked how many they had got right. They were told; 'forty' or 'fortysix' sounded a lot and they were pleased.

1.09 Instructions were paced to match each child's pace but, in order to avoid prolonging the test beyond young children's attention span, no time was wasted once the test had begun. If occasional interruption by a third party occurred the test was stopped and recontinued.

1.10 Each child appeared to work at his optimum level. Only rarely was it necessary to say something like "Think! You're in too much of a hurry. Look at the pictures carefully." In each of these cases the teacher subsequently reported such behaviour as typical of that child.

1.11 If a child spontaneously and immediately corrected his response his final answer was recorded. Which of the alternative incorrect responses the child chose was also noted on the home-made record sheet.

Scores, identity numbers, dates of birth and any other relevant information was later transferred to filing cards - one for each child. Parents' occupations were added where possible. Sources for this information were teachers, ancillary staff and occasionally school registration or admittance forms. (Usually these forms only noted the place of work for emergency contact.)

1.12 After all testing in a particular school had been completed the staff were invited to comment on the results. Generally the results matched their expectations about the children relative to each other. Fewer than a dozen of the 1928 children tested were perceived as having scored significantly less than expected. These children were then retested. Only two of them raised their scores by more than two points. One of these children had caused the examiner to entertain suspicions about the reliability of the results. As these retests were several days after the first tests the two raised scores were recorded as the true scores. Some very high scores caused surprise amongst staff not because they had not been expected to do well but because the extent of their 'superiority' had not been recognised.

1.13 Teaching and ancillary staff were helpful and friendly. Staff-room hospitality was generous and welcomed gratefully. In only one instance did the investigator sense only polite tolerance. Many headteachers seemed glad of the opportunity to discuss educational problems and innovations and to have a confidential listening ear to personal and staffing problems. Children's problems were often discussed and where relevant extra children were tested but not included in the sample.

1.14 All but one of the headteachers invited further visits and many remarked on the ease with which mutual cooperation had proceded. For the benefit of researchers seeking access to schools some headteachers' spontaneous comments on previous experiences are summarised as follows:

a Researchers sweep into school and expect me to drop everything.

b They take up too much of children's/teachers' time.

c They disrupt the day's programme.

d They come and go and never tell us much about what they are researching.

e They don't let us see the results.

f We never hear from them again. (This was the most frequent complaint)

g We are always ready to help but sometimes they leave a sour taste.

Some of the remarks may have been exaggerations and some of the 'researchers' may have been unqualified students but the comments above represent the views of many teachers.

1.15 All headteachers and staff were thanked personally and again in writing. All schools will be revisited or written to after the completion of this thesis and the final results and conclusions made available to them.

2.00 Alternate Form Equivalence

Using the same testing procedures as described above (1.00 - 1.11) 144 children were each tested with both Form A and Form B. Half of each age group were tested first with Form A and half with Form B. Within several days each child was tested with the other form. (A slightly larger number was actually tested with one form but absentees for the other form of testing were deleted from this sample.)

It was explained to the children that the investigator wanted to find out which of these two puzzles was best for children and the school children were asked later which they had liked best. (Preferences were roughly equal.)

2.01 The Sample

The children were from 14 of the 78 schools in the standardisation sample. The fourteen schools represented a range of socio-economic levels (as defined in Section i, 2.14). Table 12 shows the distribution of schools by socio-economic level and locality.

	Rural	Urban	N
all groups	2	4	6
mainly groups 1 and 2	1		1
mainly group 2		1	1
mainly groups 2 and 3	2	4	6
	5	9	14

Table 12 Distribution of schools in alternate Form subsample.

As with the standardisation sample the children were selected by birth dates (see Section i, 2.12) but children were only tested at 4, 5, 6 and 7 years. Equal numbers in each age group were, inadvertently, not quite achieved: though 36 four year olds and 36 seven year olds were tested, 34 fives and 38 sixes were tested owing to an administrative error. There were equal numbers of boys and girls in each subgroup of 18 except in the four year old group where 10 boys were tested first with Form A and 8 were tested first with Form B. When permitted, by the distribution of ages, equal numbers of boys and girls and equal numbers of each age group were tested, half in the order AB and half in the reverse order, in each single school. The first two procedures were the least feasible; the balance of the two orders of testing was approximately equal in each school.

2.02 Form B was adapted as follows:

Item 6: 'vases' replaced 'bowls' (The drawings look more like vases than bowls.)

Item 14: 'bricks' replaced 'blocks'

Item 15: '... the apple that is whole, the whole apple'

Item 17: 'Look at these trucks, lorries... the second truck.....'

Item 20: 'trailer' replaced 'wagon' (Teachers and other children were consulted about the appropriate word.)

Item 23: the wording of this item was simplified. (See Appendix 7.)

Item 27: 'has' was added at the end of the direction: '...that has as many pencils as the box has.)

Item 33: 'watch' replaced 'wristwatch'

Item 34: 'seat' replaced 'bench'

Item 35: 'shorts' replaced 'pants'

Item 37: 'middle-sized' replaced 'medium-sized' except for 7 year olds when both terms were used: '... the butterfly that is middle-sized, the medium-sized butterfly.'

Item 40: 'without any' replaced 'that has zero'

Item 46: 'miss' replaced 'skip'

Item 47: 'two' was inserted (...the two groups with equal numbers...)

Item 49: 'shop' replaced 'store'

Item 50: 'ice creams' replaced 'cones' (See Appendices 3, 4 and 7.)

Slight, non-substantive alterations were also made to the wording of items 4, 17, 19, 21 and 39. The general adaptations for individual testing that were adopted with Form A were used (see 1.02 above). The altered directions for Form B are in Appendix 7.

3.00 Predictive validity evidence

Teachers were asked to rate children who had been tested with Form A nine to twelve months previously. During the second half of the summer term 1982, letters of request, together with lists of pupils' names and rating criteria, were sent to all but one of the thirty schools where Form A testing had been completed between April and the October holiday in 1981. (The school omitted no longer existed.)

Staff were asked to provide "information on the children's present achievements preferably without referring to the (previous) results...". (A copy of the letter of request is in Appendix 9.) Three forms headed 'Spoken Language', 'Mathematics' and 'Reading' were enclosed. Copies of these are in Appendices 10, 11 and 12 respectively.

The children's names and birth dates were typed onto each form before posting. The rationale governing the design of these forms was that returns were more likely if the labour demanded of teachers was kept to a minimum and awareness of current demands on teachers' time outside normal school hours.

3.01 The Sample

Twenty nine schools were circulated with a total of 528 children's names. Twenty three sets of forms were returned with ratings of 429 children. Inevitably some children had left the district or school during the year. No forms were returned blank so presumably forms from the missing six schools were overlooked or went astray. Fourteen of the sampled schools were rural and nine were urban. The imbalance was due to the timing of the testing itinerary.

3.02 Ratings were on a three-point scale. It was thought that a finer scale would prove more problematical, with raters agonising between C and B, whereas it is comparatively easy to isolate the children who deviate from the broad middle range. This is especially true with nursery and reception children.

Teachers were asked to rate a child's current attainments as plus if he was well above average for his age group, average if he was in the broad band of average and minus if he was well below average for his age group in his school.

It was decided against asking for ratings against a national average as this might have led to confusion with some teachers under or over-estimating their pupils' against an imagined national average. Opportunity was given "to indicate that a child may be marked differently if seen against the national average... in the column headed 'Comments'." (See Appendices 10, 11 and 12.)

3.03 Criteria for rating were typed at the head of each separate form so that they were constantly under the rater's eye. The criteria, as presented to the raters, were as follows:

Spoken Language

Please disregard speech impediments and pronunciation. Take into account the child's understanding of classroom instructions, ability to frame questions, ability to maintain a conversation with children and adults, adequacy of vocabulary and the ability to describe an event coherently. If a child is rated low on language and it is clear that this is because of an emotional problem, rather than lack of ability, please indicate this in the 'Comments' column.

Mathematics A: Computational skills

Please take into account oral and written skills in counting and computation with and without concrete aids.

Mathematics B: Concepts

Concepts of space, quantity and time. Please consider the child's ability to understand and apply such concepts as behind, few, full, higher, most, smallest, wider etc. as well as practical work in measuring (linear, capacity and weight). Understanding of various time concepts such as after, next week, yesterday and the passage of time are of more relevance than clock recognition.

Reading

Please take into account the child's abilities in word recognition, reading with understanding and general reading fluency. Should this not apply to the youngest children please mark 'Pre-reader'.

4.00 Content Validity - Teachers' Talk

Eleven hours of teachers' discourse was observed in two first schools with nursery classes. The schools were not drawn in the standardisation sample and so were not used for any of the BTBC testing. They were selected because the investigator was well known and would therefore be a less threatening observer and because they drew on contrasting catchment areas. School A is in a fairly depressed mining community but occupations are by no means restricted to the pit though there are hardly any professional families living there; unemployment is rife. School B is populated mainly, but not exclusively, by the children of professional and managerial parents. Very few children from group 3 (as defined in Section i, 2.14) attend School B. All the teachers had been used to students observing them at work.

The standardisation of the test was explained to the staff without naming the BTBC; it was merely described as a language test. The teachers were asked if they would be willing to be observed in order to verify the occurrence of words in the test in everyday classroom discourse. They were told: "I shall neither be approving nor disapproving. It's no skin off my nose whether you use them or not." With this reassurance no one declined.

After all observations were completed the teachers were shown the BTBC and the completed observation sheets. Comments were invited. All but one teacher agreed that after the first five minutes they forgot the observer's presence. The exception had been obvious and the observer had left the room whenever this teacher's tension appeared to inhibit her.

4.01 In each school, observations were made in short time spans, of no longer than 30 minutes, over several days. It was thought that this would cause less tension for the teachers than would long, uninterrupted periods. It also enabled wider sampling of various kinds of classroom activities.

4.02 A stop-watch was used so that the length of duration of conversations involving the teacher could be noted. Thus the total time of eleven hours includes only the actual time that teachers were talking to or were in conversation with children.

4.03 The parcelling out of the observer's time between each teacher and various activities became much more uneven than had been planned. It was impossible to predict the precise time when staff would be engaged in an activity requiring more observation.

4.04 The distribution of the timed observations is shown below.

Age of classes	3 - 4	4 - 5	5 - 6	6 - 7	7 - 8	Total
'Activities'*	1hr 55	1hr 16	1hr 10	2hr 4	1hr 16	7hr 31
PE/movement	-	17	13	23	28	1hr 21
Class discussion	5	24	30	9	29	1hr 37
'change-over'	-	10	11	11	7	39
Total time	2hr	2hr 7	2hr 4	2hr 47	2hr 10	11hr 8
Distribution between schools	A 1hr B 1hr	A 1hr 4 B 1hr 3	A 1hr B 1hr 4	A 2hr B 47	A - B 2hr 10	A 5hr 4 B 6hr 4

Table 13 Distribution of observations between curricula activities.

* 'Activities' denotes total or partial integration of play, creative activities, mathematics, reading and writing. The imbalance of total time between the 6 and 7 year olds in the two schools is because in School A they were in two classes with older sixes and young sevens in one and a wider age range of sixes, sevens and a few eight year olds in the other.

4.05 Each time a BTBC term was spoken by a teacher it was recorded. In addition synonyms, antonyms and words related to the BTBC terms were recorded.

4.05 The limitation that no second observer was present to verify accuracy is acknowledged.

5.00 Content Validity - School Texts

The entire text of virtually every book relevant to five to eight year olds, in the following graded reading series was scrutinised for the frequency of occurrence of the BTBC terms:

Breakthrough to Literacy

Dominoes

Ladybird

Language in Action

Reading 360

Sparks

Through the Rainbow

Details of authors, publishers and titles or numbers of the books are in Appendix 13. These particular reading schemes were chosen as representing those commonly in use as advised by a teaching of reading specialist.

5.01 Similarly, pupils' workbooks from the following mathematics schemes were scrutinised:

Peak Mathematics

Mathematics for Schools (Fletcher)

Alpha Mathematics 1

Maths Adventure 1

Scottish Primary Mathematics

Advice on selection was taken from a Mathematics in Education specialist.

Books of guidance for teachers of young children were examined for the BTBC terms which featured in

a) the suggestions for teachers

b) the samples of recorded conversations with children

c) the samples of children's written work.

The books used were from the Nuffield Mathematics project and the Schools Council project "Early Mathematical Experience". Also

"Mathematics for Younger Children" (Biggs, 1971) was similarly examined. Details of all the mathematics books which were utilised are in Appendix 14.

5.02 Two books for teachers from the Schools Council project "Science 5/13" were also scrutinised page by page. These were "Early experiences: Beginnings" and "Using the environment: 1. Early explorations".

5.03 The manual for "Assessment in Nursery Education" (Bate and Smith, 1978) was scrutinised for use of the BTBC terms in questions and directions given to children.

6.00 Further content validity evidence

Observations made of teachers and children in six nurseries whose catchment areas were socially disadvantaged were scrutinised for the occurrence of the BTBC terms (Smith, 1977). These observations had been verified, at the time, by independent observers with agreements of 92%, 90% and 87% (ibid.) The observations on teachers had been made every 30 seconds but often the teacher was silent. A total of 432 observations (i.e. 216 minutes) involved a teacher's speech. These were culled from the first observation session in each nursery. (Later observations tended to be coded on the spot so that direct speech was seldom recorded.)

Recordings (by tape and by hand) of samples of 87 nursery children's speech, made during the previous investigation (ibid.), were similarly examined.

7.00 During the course of testing with the BTBC and subsequently, when the children's incorrect responses were examined, doubts arose about the adequacy of some of the drawings in the test booklets. Closely related to this concern was a desire to probe some of the children's non-linguistic strategies. Consequently, 30 children aged from 3 to 7 were each shown certain pictures extracted from the test booklets and then questioned. These children were from two schools and a day-nursery as shown in Table 14. School A is a small, village school (First-Middle) with 39 children on roll; so all the pupils present aged 5-7 were used. School B is a one-form entry First School with a nursery class; here the teachers were asked to select 'average' children who were reasonably confident and articulate. As this small investigation took place in July, most of School B's nursery children had reached their fourth birthday so two three year olds from the day-nursery were added. None of these establishments had been used for previous BTBC testing.

Age:	3		4		5		6		7		
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Total
School A:	-	-	-	1	2	2	1	2	2	3	12
School B:	2	2	4	1	1	1	1	3	1	-	16
Day-Nursery:	1	1	-	-	-	-	-	-	-	-	2
	3	3	4	2	3	3	2	4	3	3	30

Table 14 Distribution of 'picture' sample.

7.01 Materials

Each of the illustrations for Form A items 4, 24, 29, 30, 31 and 47 and Form B item 31 were cut out and mounted separately (i.e. several small cards per item). Form A items 20 and 32 were mounted intact.

7.01 Procedure

The session with each child was conducted as follows with the order of presentation unvaried:

Form A Item 4: The cards were randomly scattered (ball, bat, doll).

E: Which toy do you like best?
Tell me what they are. What's this? This? This?

Item 4 cards were removed and the picture to item 20 presented.

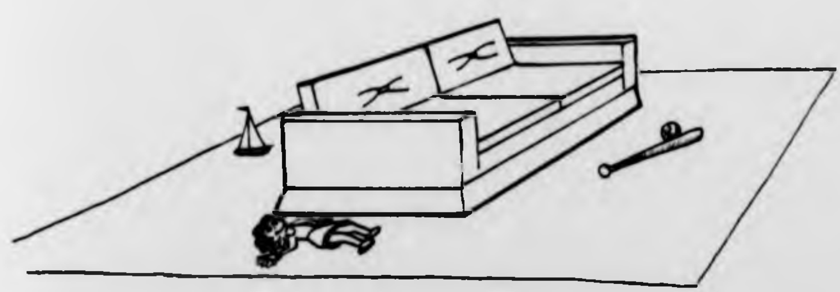


Figure 1 Form A Item 20

E: What toys can you see?

Which toy is in front of the settee?

Which toy is behind the settee?

Which toy do you like best?

Item 24: The three jars were shown one at a time in this order:

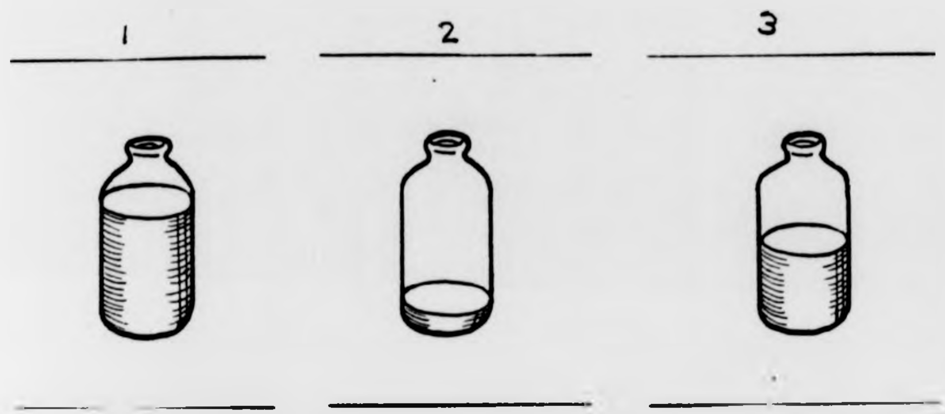


Figure 2 Form A Item 24 (reordered)

(As can be seen from Appendix 2 the standard BTBC order is 3 - 1 - 2.)

E: Has this jar got anything in it? (If 'yes') What?

(repeated for each picture)

Item 30: The cards showing an ice-cream, a pear and a piece of pie were scattered.

E: What's this? and this? and this?

Which would you like best for dinner/tea?

Are they all puddings?

Item 31: The shapes were scattered (square, triangle, pentagon, two circles).

E: Which one do you like best?

Which shapes are alike? (if appropriate) Tell me why they are alike. How are they alike?

What's that one called? (repeated three times)

Item 32: The picture of three cars was shown.

E: Which car do you like best? Why?

Item 47: The pictures of lollipops were scattered.

E: Which lollipops do you like best? Why?

Find two pictures with equal numbers of lollipops.

Item 29: The first of the squirrels was shown.

E: What is this? (If the squirrel was not named correctly the right name was supplied.)

What is the squirrel doing? And this one? And what is this one doing?

Form B Item 31: The pictures of socks were scattered.

E: Which sock do you like best?

The Method. iv. The Statistical Analysis

1.00 The standardisation of the BTBC

Even though it is doubtful that the BTBC is unidimensional this assumption was tested. The item response data obtained from the sample population were calibrated and tested for goodness of fit to the Rasch latent trait model of analysis.

1.01 The Rasch model is a one-parameter logistic model which assumes one ability parameter for each person and one difficulty parameter for each item. The model defines a general mathematical unit (logit). For ability scores, the person's ability in logits is their natural log odds for success on items of zero log odds difficulty; similarly an item's difficulty in logits is the natural log odds of failure on that item for persons of zero log odds ability. The Rasch model is consistent with 'number right' scoring. Rasch calibrated items form an item bank from which any subset of items can be drawn to make up a test (Wright, 1977).

The Rasch model is regarded by some critics as unsuitable for large-scale item banking of scholastic attainment tests (Goldstein, 1979; Goldstein and Blinkhorn, 1982; Bryce, 1981). Goldstein challenges the assumption of the model "which requires that all items appear in the same order of difficulty to all children whatever their exposure to different curricula etc." (Goldstein and Blinkhorn, 1982 p.168).

Wood (1978) criticises the rigidity of the Rasch model but poses the dilemma faced by test constructors:

"By narrowing the scope of tests in order to fit the Rasch model we may run the risk of throwing out the baby with the bath water, even though the measurements have desirable, perhaps even necessary properties, but by insisting on heterogeneous, content-rich tests, and rejecting the properties of the Rasch model, are we thereby denying ourselves the possibility of <u>any</u> coherent measurement?" (Wood, 1978 p.31).

1.02 The number of persons predicted to get an item right is not necessarily the same as the actual number who do get it right. If the observed number is identical to the expected number (i.e. if everyone behaves according to prediction) then the Rasch model is satisfied and the test items are unidimensional. If, on the other hand, there is a discrepancy between expectations and observations then clearly other variables are operating which cause the persons to respond to test items in a way that has not been predicted.

1.03 The total raw scores (Form A) obtained from the population sample were converted to <u>normalised T scores</u> (standardised scores with a mean of 50 and a standard deviation of 10). The normalisation was computed on percentiles.

2.00 <u>The internal consistency reliability</u> of the BTBC (Form A) was estimated by the application of Kuder Richardson formula 20 to total raw scores to obtain KR_{20} coefficients for each age level. The mean estimate of reliability was obtained using Fisher's Z transformation and the formula

$$Z_{av} = \frac{(N_1-9)Z_1+(N_2-9)Z_2+(N_3-9)Z_3+(N_4-9)Z_4+(N_5-9)Z_5+(N_6-9)Z_6+(N_7-9)Z_7+(N_8-9)Z_8+(N_9-9)Z_9}{(N_1-9)+(N_2-9)+(N_3-9)+(N_4-9)+(N_5-9)+(N_6-9)+(N_7-9)+(N_8-9)+(N_9-9)}$$

(McNemar, 1962). Standard errors of measurement were calculated for each age level by the formula $SE_m = SD\sqrt{1-r_{11}}$ (Anastasi, 1976).

2.01 Kuder Richardson reliability is based on the consistency of responses to all items in the test. The more homogeneous the content domain, the higher the inter-item consistency. The Kuder Richardson reliability can be conceptualised as the mean of all split-half coefficients. "Hence, unless the test items are highly homogeneous, the Kuder Richardson coefficient will be lower than the split-half reliability" (Anastasi, 1976 p.118).

3.00 Two short tests (C and D), for use with three to five year olds, were compiled by inspection of the item analysis data and observation of the children's responses during testing. The 25 items for Short Test C were selected, from Booklets 1 and 2, for their discrimination value for this age group. Short Test D consists of 20 of the 25 items in Booklet 1 (Form A). Thus Test C may be seen as the 'ideal' short version and Test D as the more 'practical' short test with the test materials as currently published.

3.01 The total raw scores of the short tests C and D were separately converted to normalised T scores. The internal consistency reliability of each short test was tested by the methods used with the whole test (see 2.00 above).

4.00 Alternate form equivalence was tested by product moment correlation of total raw scores on Form A with those on Form B. The mean correlation coefficient was calculated by use of Fisher's Z transformation as described above.

5.00 Sex differences in BTBC scores were tested by one-way analyses of variance.

6.00 Differences in BTBC scores between 'beginning-of-year' and mid-year testing were tested by one-way analyses of variance.

7.00 Construct validity

Principal components analyses were carried out on the correlation matrices of the BTBC item scores (Form A) with Varimax and Promax rotations. These analyses will provide additional evidence on the dimensionality of the BTBC.

8.00 Predictive validity was tested by product moment correlation of T scores (Form A) with teachers' ratings of pupils' attainments made one year after testing.

9.00 Content validity evidence is reported; statistical analysis was not deemed to be meaningful.

10.00 Hypotheses concerned with SES and locational differences in achievement (H_1 - H_6) were tested with one-way analyses of variance; differences between means were tested for significance with t tests.

11.00 Hypotheses concerned with semantic features and non-linguistic strategies (H_7 - H_{11}) were tested with chi-square one-sample tests where appropriate and with t tests of the difference between correlated means (McNemar, 1962).

12.00 For all tests of significance the .05 level of confidence was stipulated.

CHAPTER 9

The Results

Standardisation of the BTBC

1.00 The total number in the population sample tested with the BTBC (Form A) was 1928. Table 15 shows the distribution across nine age levels together with the mean raw score and standard deviation for each age level.

Age :	3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6
N :	115	164	193	246	228	252	232	267	231
Mean :	15.91	23.11	27.03	31.29	34.34	38.33	41.28	43.17	44.18
SD :	6.81	6.92	7.93	7.65	7.01	5.62	5.50	4.36	4.24

Table 15 Age distribution and mean raw scores

1.01 The Rasch calibration and test of goodness of fit for 50 BTBC items (Form A) produced the following results:

number of zero scores dropped = 0
number of maximum scores dropped = 30
subjects in calibration = 1898
mean = 34.69
SD = 10.21

Fit for whole test over all subjects and items:
Chi-square 71679.75 ; df 70368 p = .347

17 out of the 50 items have fit $p < .2$ whereas about 10 would be expected by chance: 14 items have fit $p < .05$, as opposed to about 2 or 3 which would be expected by chance. Therefore the Rasch latent trait model is rejected as a suitable method of analysis as the items in the BTBC are unlikely to be unidimensional.

1.02 Total raw scores (Form A) were scaled by age (N = 1928). T scores, normalised on percentiles, were computed. The T score corresponding to each raw score is shown by age level in Table 16. These are unsmoothed.

1.03 The distribution of total raw scores by age level is shown in Figures 3a - 3i on pages 206, 207, 208 and 209.

1.04 In order to obtain a clearer picture of the distribution, the total raw scores were grouped with a class interval of 5. The resultant histograms are shown in Figures 4a - 4i on pages 209, 210, 211, 212 and 213.

Raw Scores	Age 3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6
1									
2	24								
3	29								
4	31	23		21					
5									
6	33		26						
7	36								
8	38	26	29						
9	40	29		25					
10	42	32							
11	43	34	31	27					
12	45	35	33	30					
13	47	36		31			21		
14	48	37	34		24				
15	50	38	35	32					
16	51	40	36	33	27				
17	52		38	33	28				
18	53	42	39	34	29				
19	54	44	40	35	31		25		
20	56	46	41	36					
21	57	47	42	37	32	25			
22	58	48	43	37	34	28			
23	60	50	45	39	35	29	27		
24	61	51	46	41	36	31	28		
25	63	52	47	41	37	32			
26	64	54	48	43	38	32			
27	66	56	50	44	40	33			
28		57	51	45	41	34		21	
29	68	58	52	46	43	35	30	25	
30		60	53	48	44	36	31		
31	70	61	54	49	45	37	32		24
32		62	55	50	46	38	34	28	28
33	74	63	57	51	47	39	36	31	31
34		65	58	52	49	41	37	33	32
35		67	59	53	50	43	37	34	33
36		69	62	55	52	44	39	35	34
37		71	63	57	53	46	41	37	35
38		74	64	59	54	48	43	38	37
39		77	66	61	55	50	44	40	38
40			69	63	57	52	46	42	40
41			72	65	59	54	48	41	41
42			76	67	62	56	50	46	43
43				68	64	58	52	48	45
44				70	65	60	54	50	48
45				73	66	63	57	53	50
46					68	66	59	56	53
47					71	69	61	59	55
48					74	73	65	62	59
49				79	79	75	69	65	63
50						79	72	70	69

(N = 1928)

Table 16 T scores with corresponding raw scores by age level (unsmoothed). BTBC Form A.

Fig. 3a 3:6

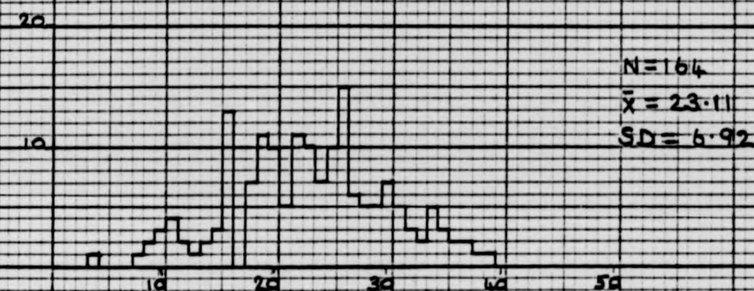


Fig. 3b 4:0

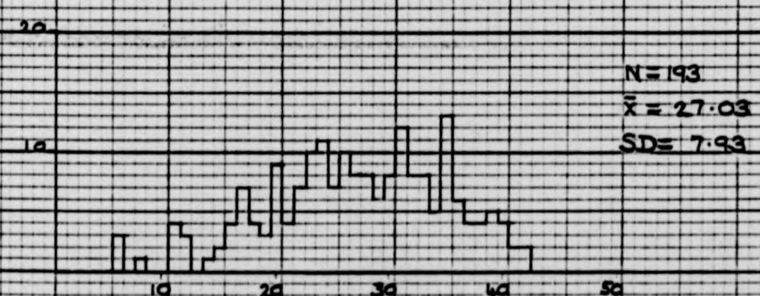


Fig. 3c 4:6

Fig. 3 Distribution of raw scores by age level.

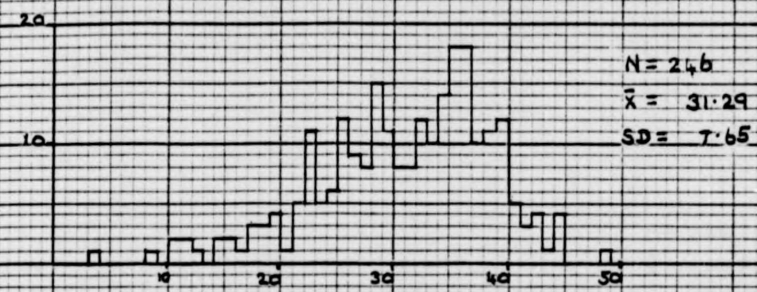


Fig. 3 d 5:0

Fig. 3 e 5:6

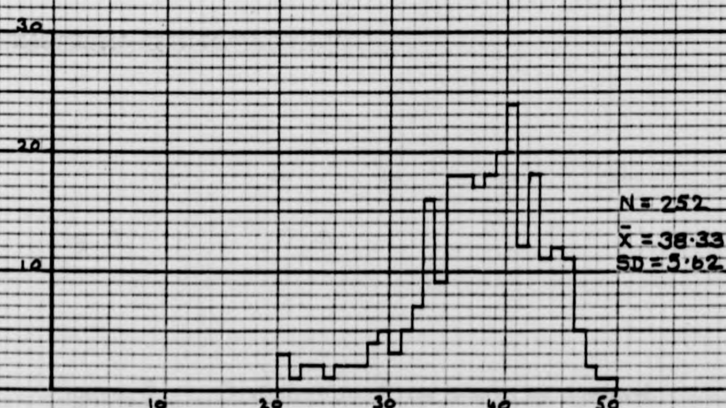


Fig. 3 f 6:0

Figure 3 Distribution of raw scores by age level.

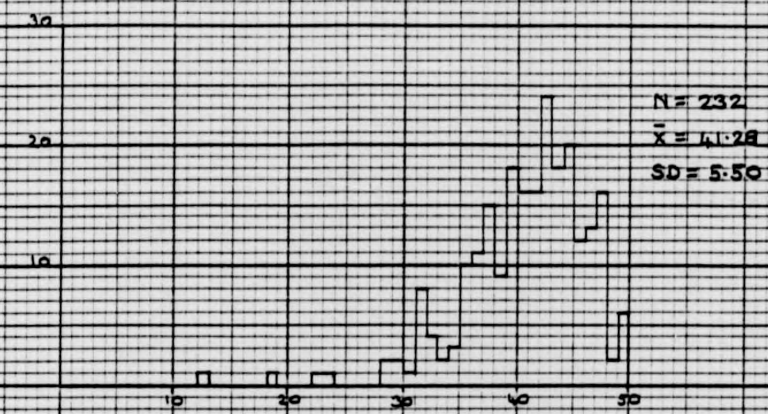


Fig. 3 g   6:6

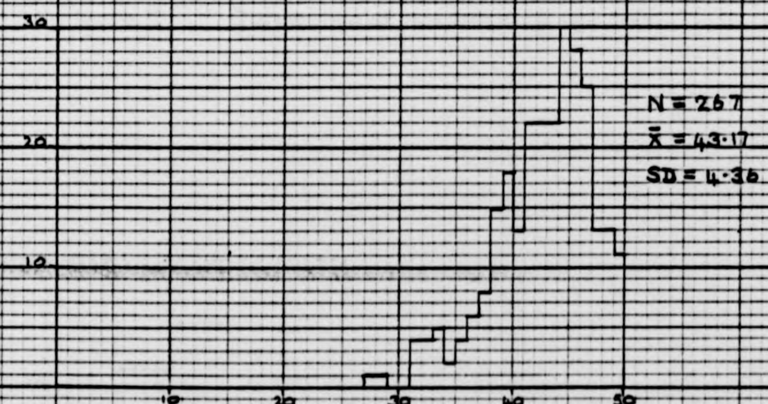


Fig. 3 h   7:0

Figure 3   Distribution of raw scores by age level.

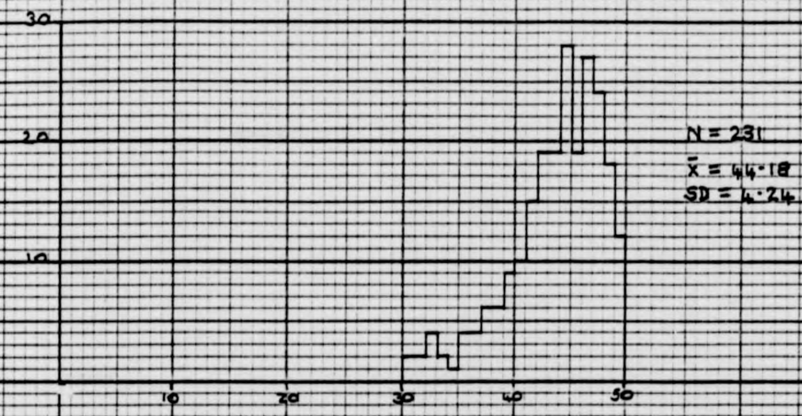


Fig. 3 i 7:6

Figure 3 Distribution of raw scores by age level.

Fig. 4 a 3:6

Figure 4 Distribution of raw scores by age level.

Scores grouped with a class interval of 5.

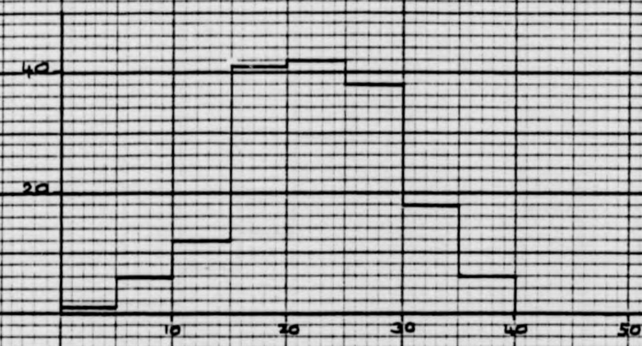


Fig. 4 b 4:0

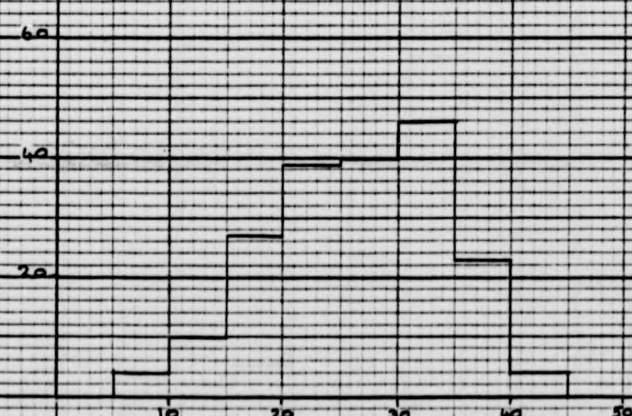


Fig. 4 c 4:6

Figure 4 Distribution of raw scores by age level.

Scores grouped with a class interval of 5.

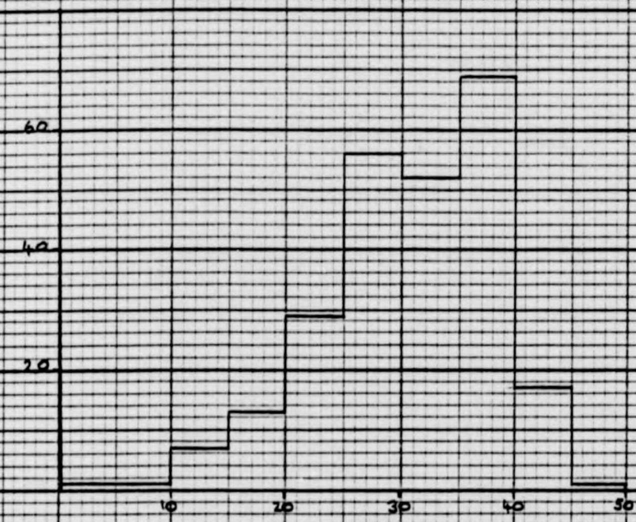


Fig. 4 d 5:0

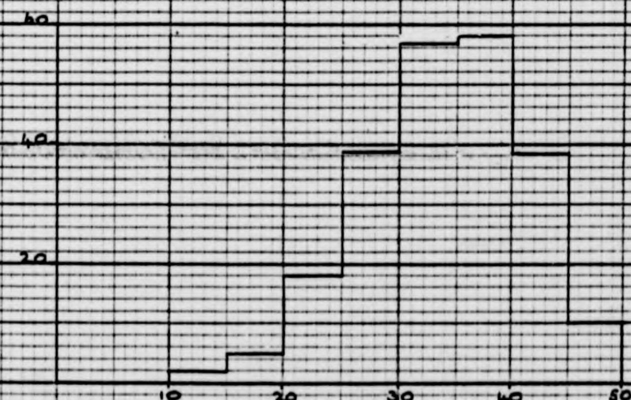


Fig. 4 e 5:6

Figure 4 Distribution of raw scores by age level.

Scores grouped with a class interval of 5.

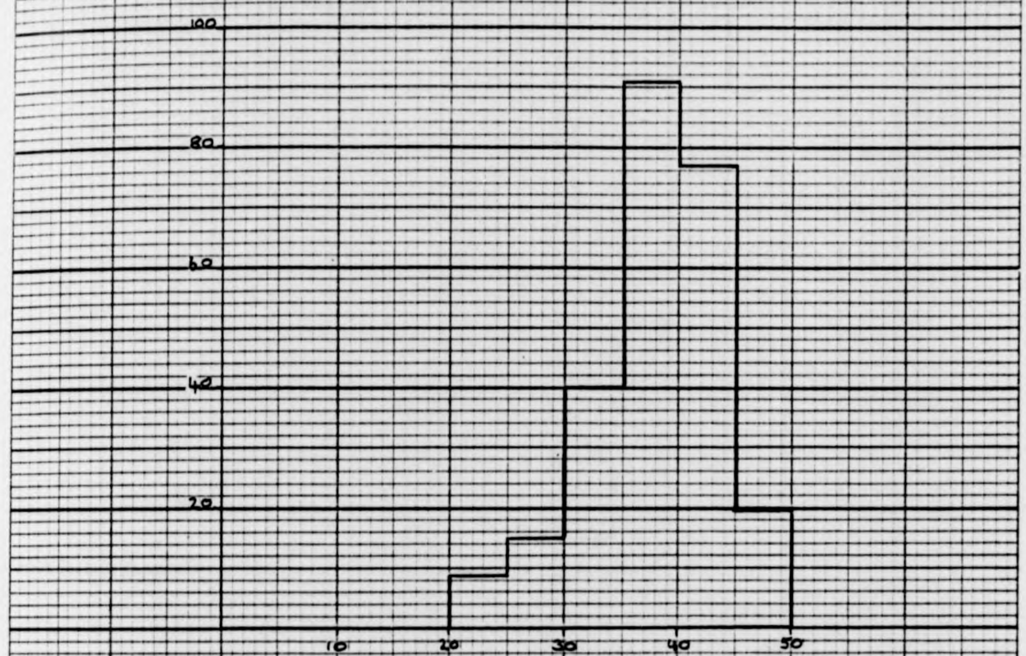


Fig. 4 f 6:0

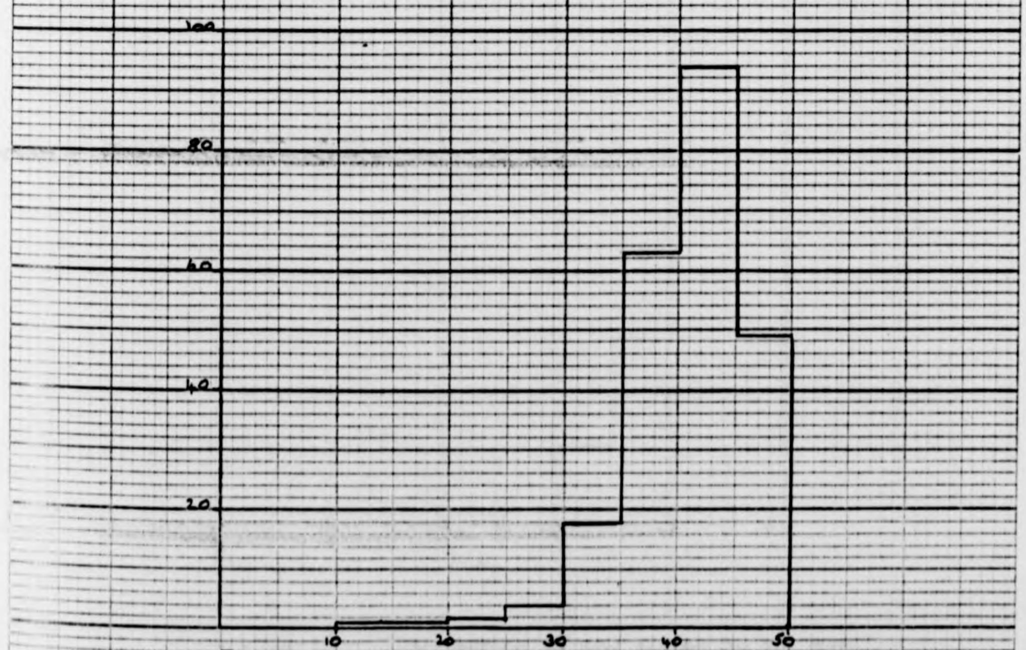


Fig. 4 g 6:6

Figure 4 Distribution of raw scores by age level.

Scores grouped with a class interval of 5.

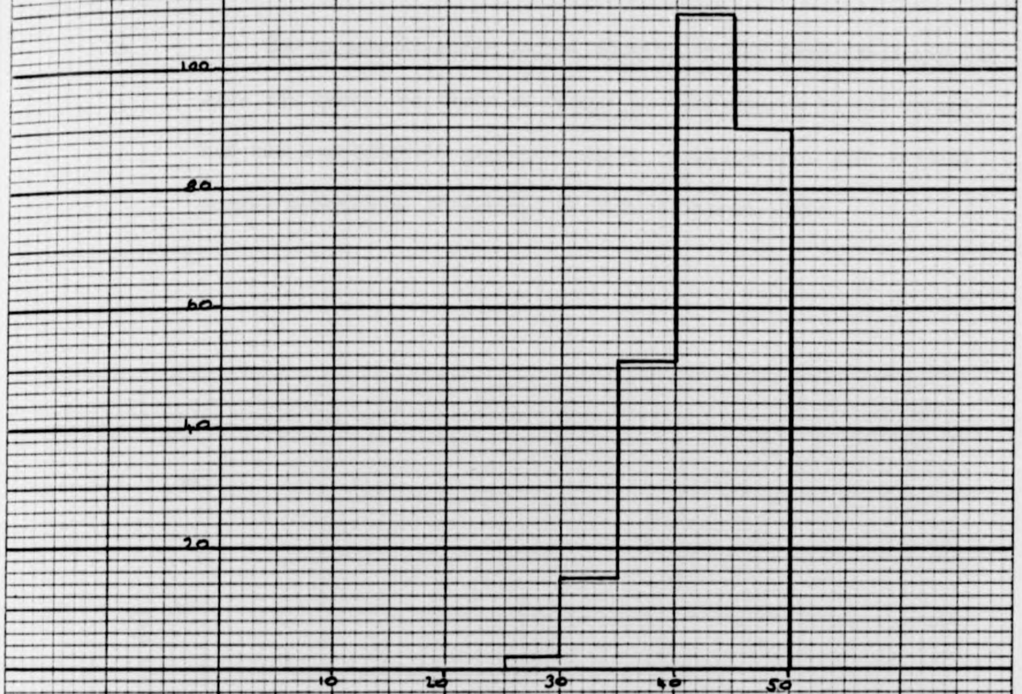


Fig. 4 h 7:0

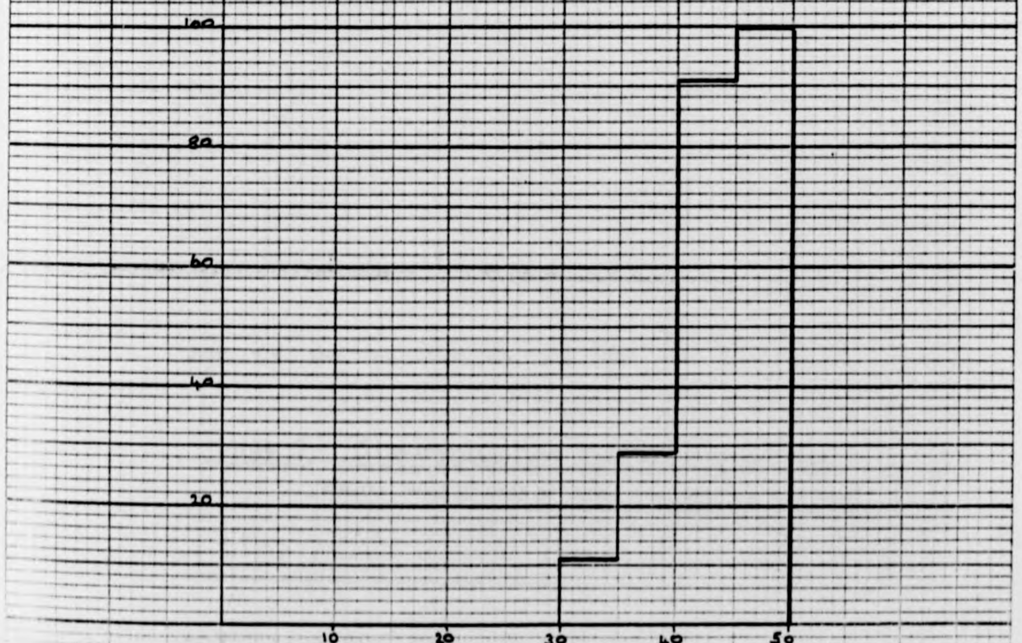


Fig. 4 i 7:6

Figure 4 Distribution of raw scores by age level.

Scores grouped with a class interval of 5

1.05 The standardisation curve, plotted by the mean raw score for each age level, is shown in Figure 5. This curve is not smooth but the acceleration around 4 years, coincidental with several months experience of nursery or school, makes sense.

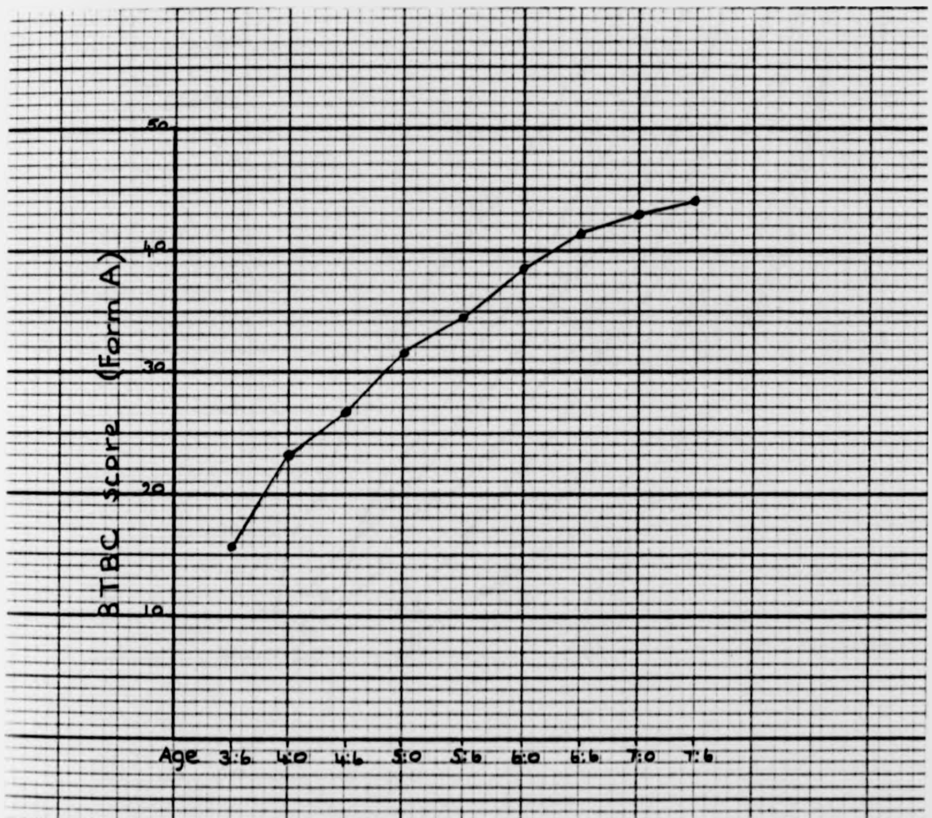


Figure 5 Standardisation curve (unsmoothed): Mean raw scores at each age level. (N = 1928)

1.06 Development curves for each age level plotted by raw scores against normalised T scores are shown in Figure 6.

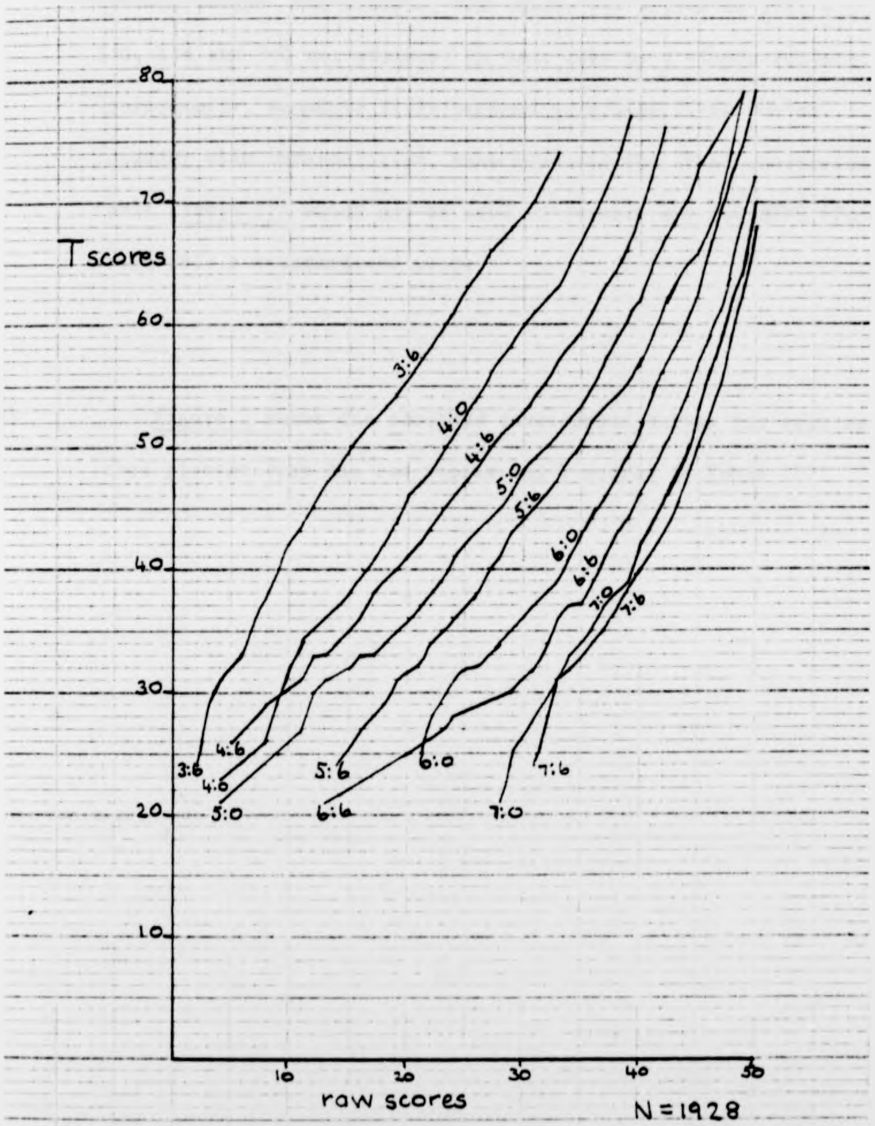


Figure 6 Development curves (unsmoothed).

1.07 If Fig. 6 is inspected in conjunction with Figs. 3a - 3i it can be seen that some very small numbers of subjects have produced anomalies in the T scores, especially at the lower end of the scoring range. For example, the lowest raw scores at 3:6, 4:0, 5:0, 6:6 and 7:0 are produced in each case by a single subject. Consequently, smoother distributions have been fitted which, together with interpolation, have produced the final conversion table (Table 17) which may be used to convert any raw BTBC (Form A) score into a standardised score.

1.08 Figure 7 shows the smoothed development curves for each age level plotted from the raw scores and T scores in Table 17.

Raw Score	Age 3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6
1									
2	26								
3	28								
4	29	26	25	20					
5	31	27	26	21					
6	33	29	27	22					
7	36	30	28	23					
8	38	31	29	24					
9	40	32	30	25					
10	42	33	31	26					
11	43	34	32	27					
12	45	35	33	28					
13	47	36	34	29	24	19	14		
14	48	37	35	30	25	20	15		
15	50	38	36	31	26	21	16		
16	51	40	37	32	27	22	17		
17	52	41	38	33	28	23	18		
18	53	42	39	34	29	24	19		
19	54	44	40	35	30	25	20		
20	56	46	41	36	31	26	21		
21	57	47	42	37	32	27	22		
22	58	48	43	38	34	28	23		
23	60	50	45	39	35	29	24		
24	61	51	46	41	36	30	25		
25	63	52	47	42	37	31	26		
26	64	54	48	43	38	32	27		
27	66	56	50	44	40	33	28		
28	67	57	51	45	41	34	29	25	
29	68	58	52	46	42	35	30	26	
30	69	60	53	48	44	36	31	27	
31	70	61	54	49	45	37	32	28	24
32	72	62	55	50	46	38	34	29	25
33	74	63	57	51	47	39	36	30	26
34		65	58	52	49	41	37	31	28
35		67	59	53	50	43	38	32	30
36		69	61	55	52	44	39	34	32
37		71	62	57	53	46	41	36	34
38		74	64	59	54	48	43	38	36
39		77	66	61	55	50	44	40	37
40			69	63	57	52	46	42	39
41			72	65	59	54	48	44	41
42			76	67	61	56	50	46	43
43				68	63	58	52	48	45
44				70	65	60	54	50	48
45				73	67	63	57	53	50
46				76	69	66	59	56	53
47				78	72	69	61	59	56
48				80	75	73	65	62	59
49				82	79	75	69	65	63
50						79	72	70	68

Table 17 Conversion of raw scores to standard scores

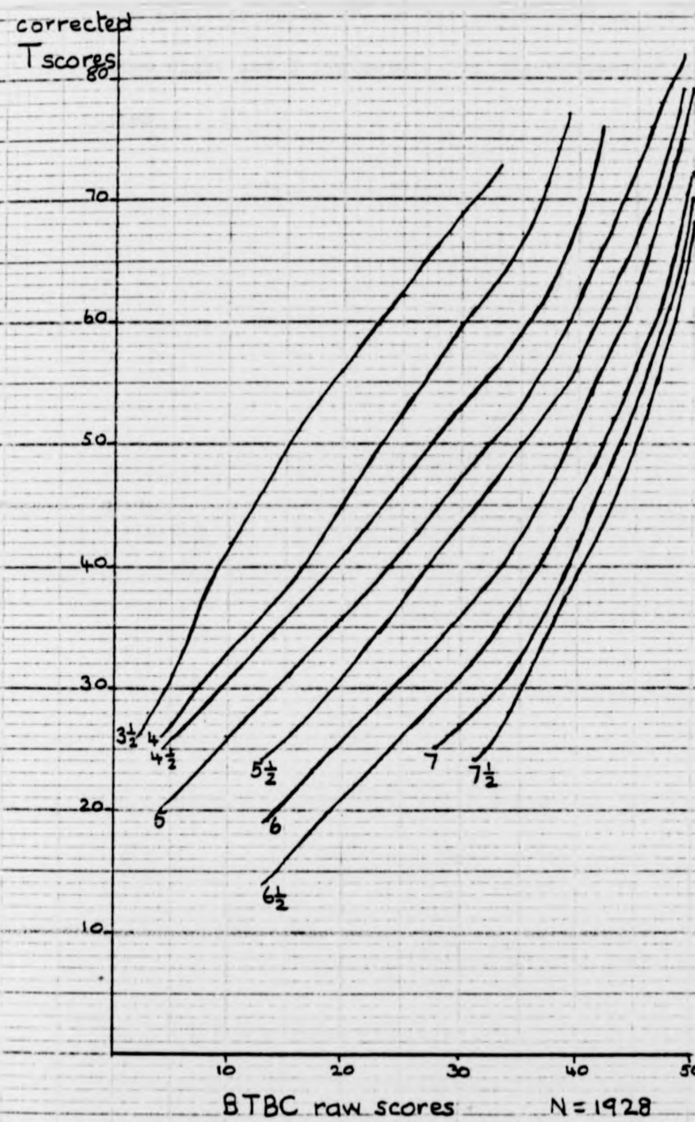


Figure 7 Developmental curves after smoothing.

1.09 The distribution of item responses is shown in Table 18. The facility value (F) is the proportion of subjects getting the item right and D represents the discrimination value for each item.

	3:6		4:0		4:6		5:0		5:6		6:0		6:6		7:0		7:6	
Item	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D
1	.69	.60	.85	.24	.92	.18	.96	.11	.96	.11	.99	.04	1.00	.00	1.00	.01	1.00	.00
2	.66	.44	.82	.28	.92	.19	.94	.10	.98	.05	.99	.01	1.00	.00	1.00	.00	1.00	.00
3	.41	.38	.55	.52	.65	.53	.74	.41	.84	.29	.92	.10	.92	.13	.97	.04	.98	.02
4	.71	.40	.84	.32	.91	.18	.93	.15	.98	.04	.98	.06	1.00	.01	1.00	.00	.99	.03
5	.83	.20	.86	.18	.90	.13	.94	.13	.93	.10	.98	.05	.98	.10	.99	.01	1.00	.02
6	.38	.41	.62	.37	.77	.45	.85	.31	.87	.16	.93	.15	.98	.04	.94	.08	.98	.08
7	.32	.64	.63	.49	.78	.42	.88	.26	.93	.14	.98	.02	.99	.03	1.00	.00	1.00	.02
8	.46	.44	.68	.37	.76	.27	.81	.32	.81	.03	.81	.15	.76	.22	.79	.14	.85	.14
9	.33	.64	.60	.51	.77	.50	.86	.35	.90	.10	.92	.18	.97	.07	1.00	.01	.97	.06
10	.59	.18	.79	.34	.85	.24	.92	.16	.97	.10	.98	.06	.99	.03	1.00	.01	1.00	.02
11	.74	.44	.83	.07	.91	.18	.92	.13	.97	.03	.98	.04	.99	.03	1.00	.01	1.00	.00
12	.37	.22	.57	.44	.67	.30	.74	.22	.75	.26	.88	.13	.89	.20	.89	.14	.94	.05
13	.47	.57	.65	.39	.68	.52	.82	.29	.92	.12	.97	.05	1.00	.01	1.00	.00	1.00	.02
14	.18	.39	.42	.35	.58	.58	.73	.54	.73	.43	.82	.27	.92	.11	.90	.16	.91	.17
15	.33	.22	.44	.36	.51	.51	.65	.47	.72	.45	.86	.25	.94	.14	.96	.12	.97	.06
16	.91	.27	.92	.10	.98	.03	.98	.05	1.00	.00	1.00	-.01	.99	.01	1.00	.00	1.00	.00
17	.25	.19	.26	.11	.33	.40	.55	.54	.71	.52	.86	.33	.90	.22	.95	.08	.97	.02
18	.28	.28	.45	.54	.59	.58	.73	.40	.81	.26	.94	.12	.93	.13	.98	.05	.97	.08
19	.42	.61	.67	.32	.65	.45	.72	.23	.72	.15	.69	.07	.72	.12	.76	.20	.78	.25
20	.44	.51	.58	.41	.68	.39	.74	.38	.77	.31	.91	.21	.94	.13	.95	.09	.94	.06
21	.44	.71	.59	.57	.76	.45	.74	.55	.86	.32	.95	.15	.94	.10	.99	.04	1.00	.02
22	.30	.15	.42	.40	.58	.48	.68	.43	.69	.46	.86	.28	.86	.24	.93	.16	.93	.20
23	.10	.19	.28	.14	.36	.32	.49	.37	.58	.30	.63	.29	.65	.50	.74	.40	.79	.27
24	.76	.37	.92	.05	.90	.11	.94	.08	.94	.03	.95	-.01	.96	.03	.97	.00	.97	.02
25	.28	.22	.40	.15	.46	.01	.52	.22	.57	.29	.63	.40	.77	.32	.88	.20	.84	.24
26	.29	.42	.45	.07	.44	.01	.39	-.03	.47	.14	.42	.29	.51	.33	.55	.44	.65	.42
27	.08	.19	.32	.37	.57	.64	.73	.50	.87	.25	.94	.16	.96	.11	.98	-.01	.96	.11
28	.14	.42	.30	.30	.45	.45	.42	.33	.49	.42	.62	.33	.60	.36	.65	.33	.68	.42
29	.19	.35	.32	.30	.30	.31	.44	.29	.50	.57	.59	.33	.72	.35	.70	.38	.76	.26
30	.18	.23	.40	.35	.50	.58	.64	.54	.70	.47	.87	.20	.93	.17	.93	.08	.92	.23

<u>Table 18</u>

continued over...

	3:6		4:0		4:6		5:0		5:6		6:0		6:6		7:0		7:6	
Item	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D
31	.26	.26	.37	.18	.26	.28	.39	.34	.39	.39	.58	.38	.63	.59	.72	.42	.76	.43
32	.10	.23	.21	.26	.29	.58	.48	.62	.60	.47	.78	.45	.88	.27	.91	.20	.91	.20
33	.06	.10	.20	.31	.30	.44	.43	.55	.50	.48	.64	.48	.73	.45	.79	.24	.84	.31
34	.49	.44	.71	.27	.77	.40	.88	.22	.88	.25	.96	.10	.97	.09	.99	.04	.98	.06
35	.28	.48	.34	.19	.40	.37	.52	.50	.59	.36	.64	.44	.69	.28	.80	.23	.81	.34
36	.17	.42	.35	.32	.43	.58	.39	.38	.51	.33	.56	.31	.66	.29	.69	.32	.74	.25
37	.16	.16	.31	.17	.28	.36	.45	.44	.51	.53	.62	.56	.77	.49	.77	.53	.80	.48
38	.18	.48	.30	.27	.43	.35	.45	.42	.55	.25	.64	.24	.74	.25	.70	.29	.78	.29
39	.07	.16	.22	.34	.42	.59	.49	.56	.54	.61	.70	.57	.78	.61	.85	.37	.92	.26
40	.45	.67	.77	.43	.84	.39	.92	.14	.98	.04	.99	.02	.99	.04	1.00	.01	1.00	.00
41	.41	.54	.52	.52	.63	.45	.70	.50	.75	.36	.86	.12	.85	.34	.93	.21	.94	.14
42	.22	.35	.40	.40	.51	.56	.66	.46	.70	.52	.83	.34	.93	.14	.93	.17	.95	.14
43	.23	.52	.29	.09	.33	.44	.40	.35	.50	.35	.58	.39	.69	.48	.84	.35	.87	.30
44	.15	.39	.28	.20	.36	.41	.41	.14	.44	.33	.48	.50	.60	.50	.64	.41	.76	.30
45	.04	.10	.14	.16	.15	.20	.18	.23	.19	.24	.29	.46	.38	.48	.47	.63	.55	.62
46	.01	.03	.05	.10	.09	.13	.24	.37	.40	.60	.57	.49	.75	.52	.87	.26	.89	.22
47	.04	.10	.06	.13	.11	.26	.28	.47	.43	.46	.56	.43	.73	.38	.72	.42	.75	.40
48	.02	.07	.04	.08	.13	.30	.26	.35	.35	.30	.43	.47	.64	.53	.76	.45	.79	.48
49	.01	.03	.05	.11	.11	.31	.17	.34	.33	.63	.47	.47	.67	.45	.79	.31	.81	.37
50	.04	.07	.12	.16	.12	.12	.20	.14	.26	.34	.35	.49	.53	.54	.65	.64	.64	.64

...continued from page 220

Table 18 Distribution of item responses by age level.

2.00 Reliability estimates

The Kuder Richardson (Formula 20) internal consistency coefficients and the standard errors of measurement obtained for the BTBC (Form A) total raw scores at each age level are shown in Table 19 together with the mean reliability estimates and the mean estimate of the standard error of measurement.

	Age									
	3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6	Mean
KR_{20}	.82	.80	.85	.85	.83	.78	.80	.74	.75	.81
SE_m	2.9	3.1	3.0	3.0	2.9	2.6	2.5	2.2	2.1	2.7
N	115	164	193	246	228	252	232	267	231	

Table 19 Reliability estimates and standard errors of measurement.

3.00 The 25 items selected for the short test C from Form A booklets 1 and 2 are shown in Table 20 together with their discrimination values at the younger age levels.

BTBC items		Discrimination values			
	Age:	3:6	4:0	4:6	5:0
Booklet 1					
1	top	.603	.235	.117	.111
2	through	.438	.284	.194	.101
3	away from	.381	.524	.531	.413
4	next to	.402	.318	.177	.150
6	some, not many	.414	.374	.435	.309
7	middle	.643	.489	.419	.261
9	farthest	.643	.506	.499	.347
12	widest	.215	.440	.304	.223
13	most	.574	.389	.515	.289
14	between	.387	.346	.578	.535
15	whole	.218	.363	.512	.466
18	corner	.286	.542	.579	.401
20	behind	.508	.408	.385	.380
21	row	.706	.571	.451	.546
22	different	.152	.395	.481	.427
Booklet 2					
27	as many	.194	.366	.644	.497
28	side	.419	.302	.446	.326
30	other	.225	.347	.577	.538
32	not the first or the last	.226	.255	.575	.622
33	never	.097	.305	.444	.551
34	below	.443	.273	.402	.224
36	always	.419	.317	.576	.376
40	without any	.672	.434	.387	.138
41	above	.544	.524	.448	.500
42	every	.353	.396	.562	.456

No. of items = 25

Table 20 Short test C items.

3.01 The 20 items selected from Form A booklet 1 for the short test D are shown in Table 21 together with their discrimination values.

BTBC items	Discrimination values			
Age:	3:6	4:0	4:6	5:0
Booklet 1				
1 top	.603	.235	.177	.111
2 through	.438	.284	.194	.101
3 away from	.381	.524	.531	.413
4 next to	.402	.318	.177	.150
5 inside	.204	.184	.128	.125
6 some, not many	.414	.374	.435	.309
7 middle	.643	.489	.419	.261
8 few	.443	.371	.272	.324
9 farthest	.643	.506	.499	.347
10 around	.175	.335	.241	.162
11 over	.437	.072	.177	.127
12 widest	.215	.440	.304	.223
13 most	.574	.389	.515	.289
14 between	.387	.346	.578	.535
15 whole	.218	.363	.512	.466
17 second	.191	.105	.395	.544
18 corner	.286	.542	.579	.401
20 behind	.508	.408	.385	.380
21 row	.706	.571	.451	.546
22 different	.152	.395	.481	.427

No. of items = 20

Table 21 Short test D items

3.02 The distribution of total raw scores by age level are shown in Figures 8a - 8d for Short Test C and in Figures 9a - 9d for Short Test D.

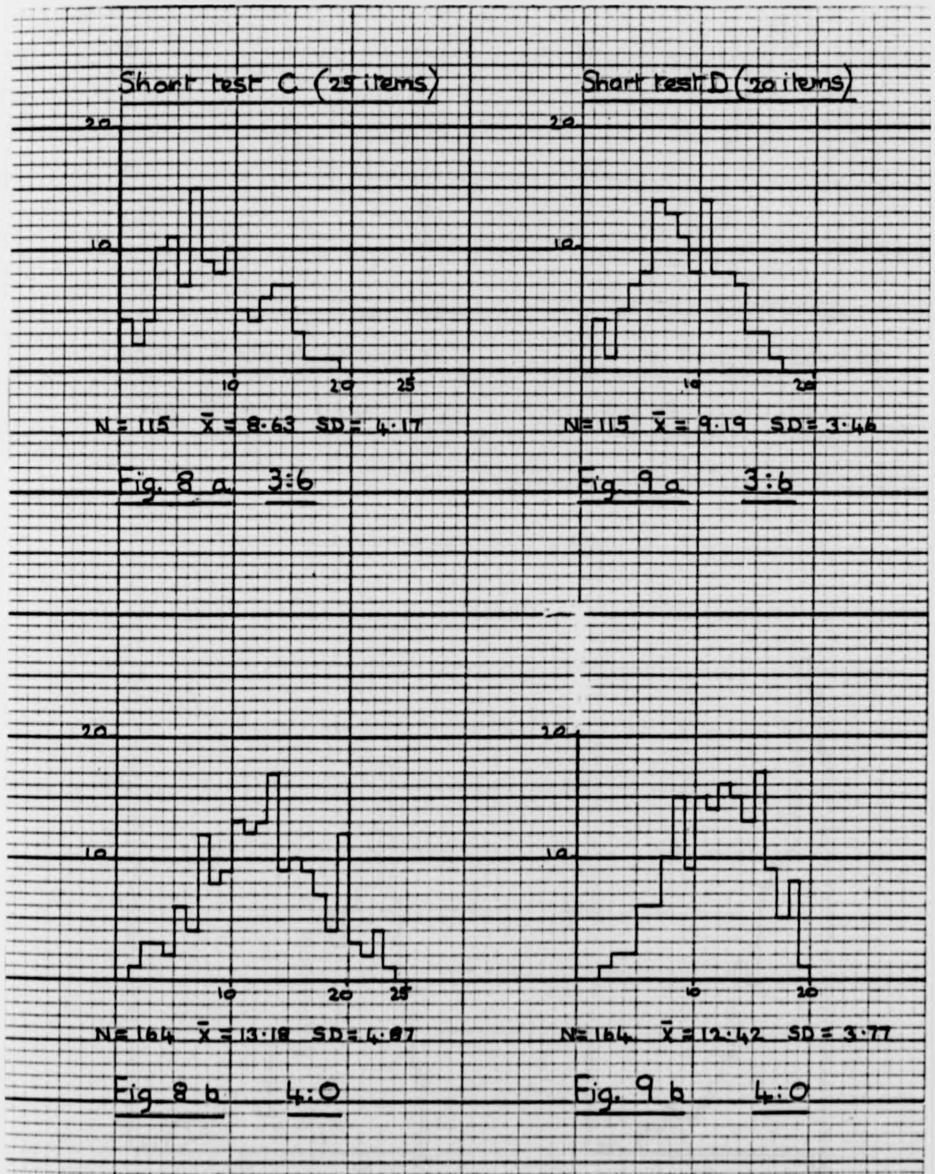


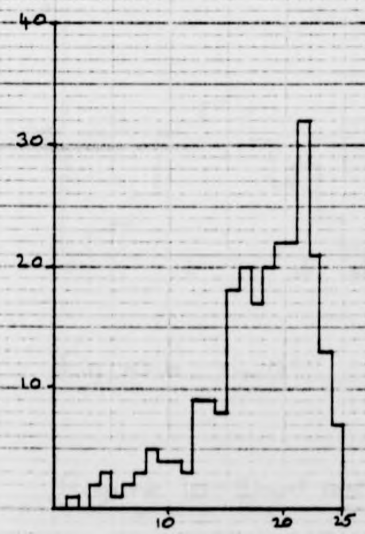
Figure 8 Distribution of raw scores by age level. Short test C.

Figure 9 Distribution of raw scores by age level. Short test D.

Short test C (25 items)

N= 193 $\bar{x}$ = 16·04 SD = 5·22

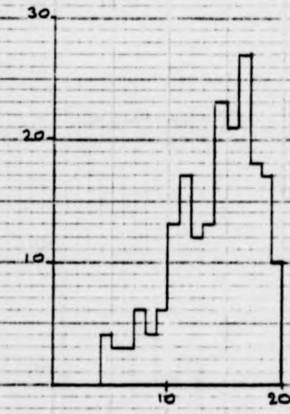
Fig. 8 c 4:6



N= 246 $\bar{x}$ = 18·24 SD = 4·73

Fig. 8 d 5:0

Short test D (20 items)



N= 193 $\bar{x}$ = 14·51 SD = 3·65

Fig. 9 c 4:6

Fig. 9 d 5:0

3.03 The total raw scores were grouped with a class interval of 2. The resultant histograms are shown in Figures 10a - 10d for Short Test C and in Figures 11a - 11d for Short Test D.

Short test C

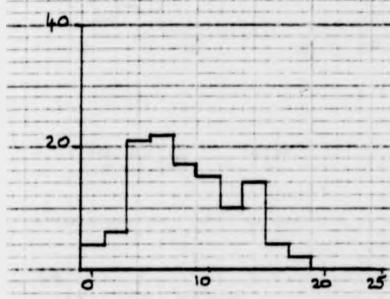


Fig. 10 a 3:6

Short test D

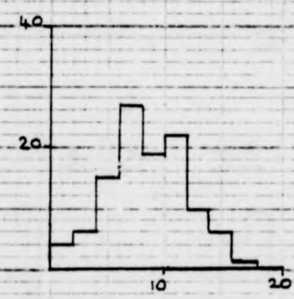


Fig. 11 a 3:6

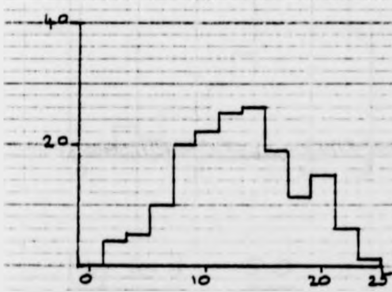


Fig. 10 b 4:0

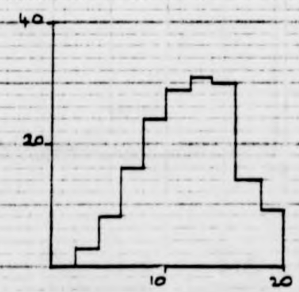


Fig. 11 b 4:0

Figure 10 Short test C. Figure 11 Short test D.

Distribution of raw scores by age level.

Scores grouped with a class interval of 2.

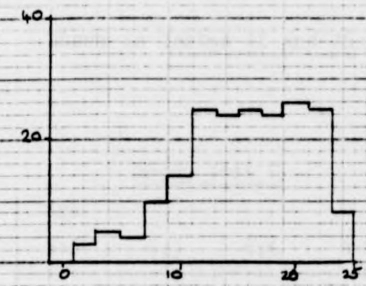


Fig. 10 c 4:6

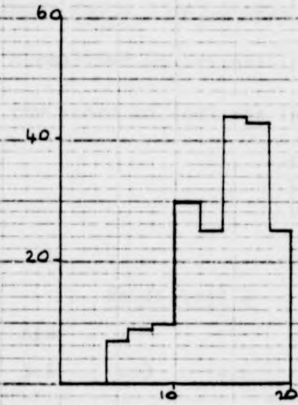


Fig. 11 c 4:6

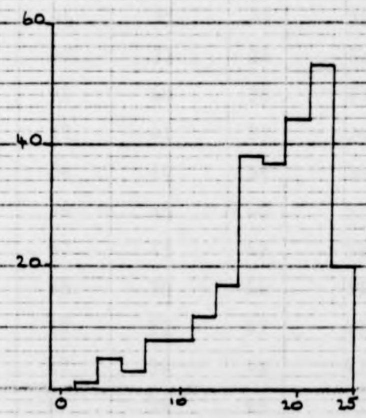


Fig. 10 d 5:0

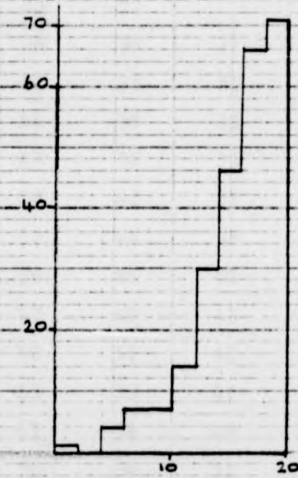


Fig. 11 d 5:0

Figure 10 Short test C.

Figure 11 Short test D.

Distribution of raw scores by age level.

Scores grouped with a class interval of 2.

3.04 The total raw scores of the short tests C and D were separately converted to T scores, normalised on percentiles. The T score corresponding to each raw score at each age level is shown in Tables 22 and 23.

Raw Score	Short test C Age 3:6	4:0	4:6	5:0	Short test D Age 3:6	4:0	4:6	5:0
1	29							
2	33	23	24	21	29			21
3	35	28	28		32	23		
4	39	32	31	26	35	27		
5	42	33	32	29	38	30	27	26
6	45	35	33	31	41	33	31	29
7	47	37	34	32	44	36	33	31
8	50	40	36	33	47	39	35	33
9	52	42	37	34	50	42	37	34
10	54	44	39	36	52	44	38	35
11	56	46	40	37	54	46	40	37
12	57	48	42	38	57	49	43	38
13	58	50	44	39	60	51	45	40
14	60	52	47	40	63	53	47	43
15	63	54	48	42	67	56	49	45
16	67	56	49	43	70	59	52	47
17	70	57	51	46	76	62	56	50
18	72	59	53	48		65	60	53
19	76	61	54	50		69	63	58
20		63	56	52		77	69	65
21		66	58	54				
22		68	61	57				
23		71	64	62				
24		77	70	66				
25				72				

N = 718

Table 23 Short Test D: T scores (unsmoothed) with corresponding raw scores.

N = 718

Table 22 Short Test C: T scores (unsmoothed) with corresponding raw scores.

3.05 The standardisation curves for the short tests are shown in Figure 12.

3.06 Development curves for each age level (3:6 - 5:0) plotted by raw scores against the normalised T scores for the short tests are shown in Figure 13.

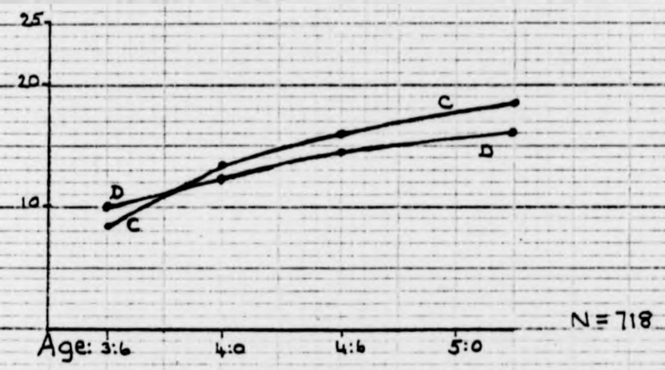


Fig. 12. Standardisation curves for tests C and D: mean raw scores at each age level.

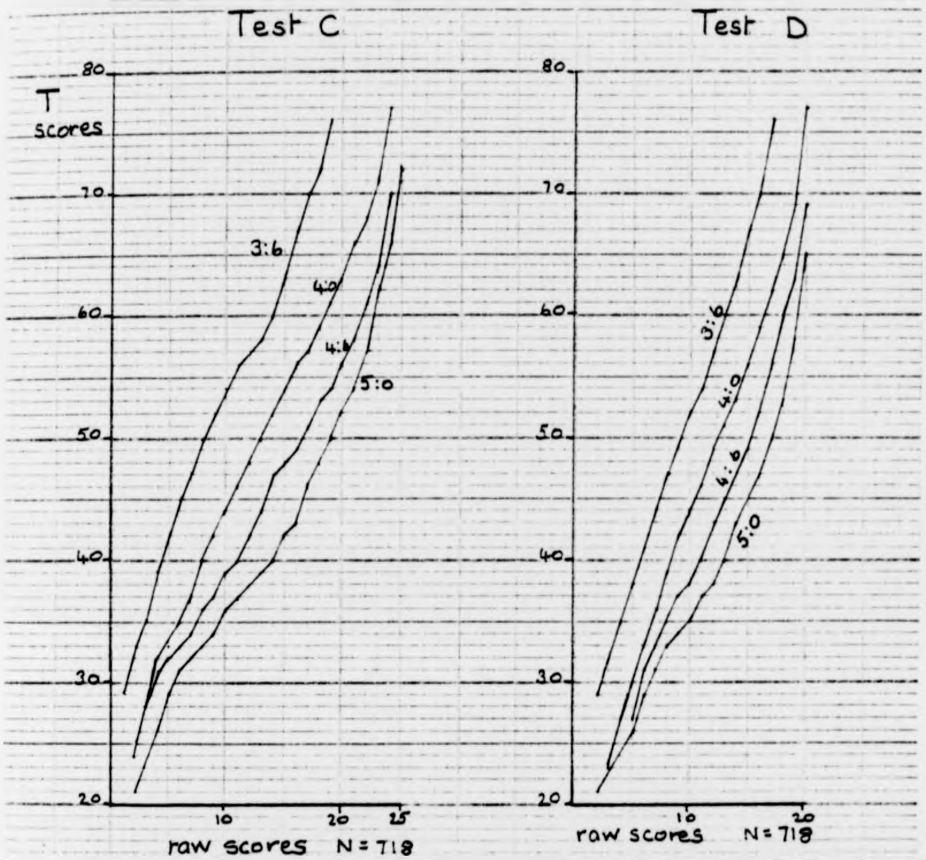


Figure 13 Tests C and D: developmental curves (unsmoothed).

3.07 The final conversion tables for the short tests C (Table 24) and D (Table 25) were arrived at after interpolation and smoothing as described earlier.

Raw Score	Short test C Age 3:6	4:0	4:6	5:0	Short test D Age 3:6	4:0	4:6	5:0
1	29							
2	33	27	24	21	29	23	22	20
3	35	29	26	23	32	25	24	22
4	39	31	28	25	35	28	26	24
5	42	33	30	28	38	30	28	26
6	45	35	32	30	41	33	30	28
7	47	37	34	32	44	36	32	30
8	50	40	36	33	47	39	34	32
9	52	42	37	34	50	42	36	34
10	53	44	39	35	52	44	38	35
11	55	46	40	37	54	46	40	37
12	57	48	42	38	57	49	43	39
13	58	50	44	39	60	51	45	41
14	60	52	46	40	63	53	47	43
15	63	54	48	42	67	56	49	45
16	66	56	49	44	70	59	52	47
17	69	57	51	46	76	62	56	50
18	72	59	53	48		65	60	53
19	76	61	54	50		69	64	58
20		63	56	52		77	69	65
21		66	58	54				
22		68	61	57				
23		71	64	61				
24		77	70	66				
25				72				

Table 24 Test C: conversion of raw scores to standard scores.

Table 25 Test D: conversion of raw scores to standard scores.

3.08 Figure 14 shows the smoothed development curves for each age level plotted from the raw scores and T scores in the final conversion tables for Tests C and D (Tables 24 and 25).

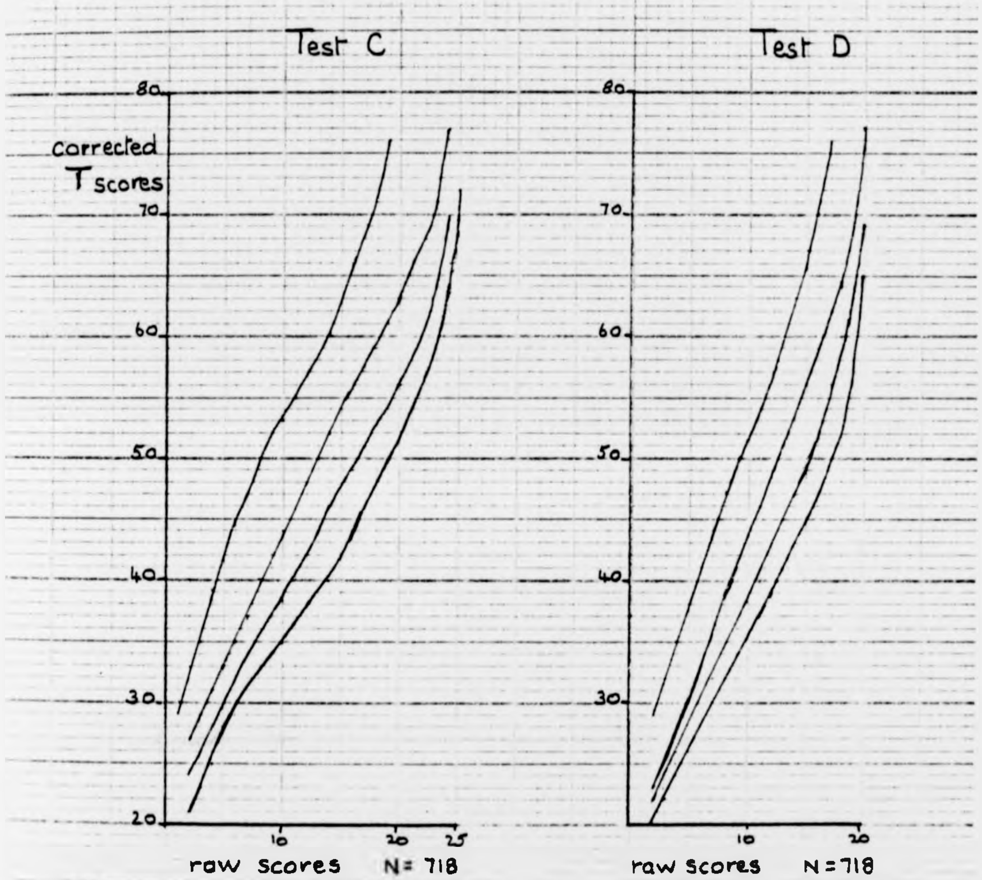


Figure 14 Tests C and D: development curves (smoothed).

3.09 Reliability estimates of Tests C and D

Table 26 shows the KR_{20} coefficients and the standard errors of measurement for the total raw scores of the short tests C and D at each age level together with the mean reliability estimate and the mean standard error of measurement.

	Age 3:6	4:0	4:6	5:0	Mean
Test C					
Mean	8.69	13.18	16.04	18.24	
SD	4.17	4.87	5.22	4.73	
KR_{20}	.72	.77	.82	.81	.79
SE_m	2.2	2.3	2.2	2.1	2.2
Test D					
Mean	9.19	12.42	14.68	16.13	
SD	3.46	3.77	3.77	3.50	
KR_{20}	.64	.71	.75	.77	.74
SE_m	2.1	2.0	1.9	1.7	1.9
N	115	164	193	246	Total: 718

Table 26 Reliability estimates of Tests C and D.

3.10 Table 27 shows the correlation matrix of the T scores (unsmoothed) obtained by 718 children on the short tests C and D and on the whole test. As tests C and D are each subscores of the whole test and as 17 items are common to both Tests C and D the correlation coefficients are bound to be spuriously high (McNemar, 1962).

	C	D	BTBC (50 items)
Test C (25 items)		.910	.935
Test D (20 items)			.889
Mean	50.02	49.88	50.00
SD	9.83	9.73	9.97

Table 27 Correlation matrix of Tests C, D and Form A.

4.00 Sex differences

Table 28 shows the results of one-way analyses of variance between the mean scaled BTBC scores of the total sample of boys and girls.

Total sample	Boys	Girls
N	965	963
Mean T score	49.71	50.29
SD	10.15	9.65
F	1.68 df 1927; NS	

Table 28 ANOVA: T scores of boys and girls

4.01 The null hypothesis that there is no significant difference between boys' and girls' mean scaled BTBC scores is accepted.

4.02 Table 29 summarises the mean scores at each age level.

Age	N		Mean T score		SD	
	Boys	Girls	Boys	Girls	Boys	Girls
3:6	57	58	49.42	50.43	10.46	9.64
4:0	84	80	50.58	49.33	10.14	9.80
4:6	96	97	49.02	51.07	9.70	10.29
5:0	114	132	49.32	50.64	10.36	9.71
5:6	115	113	48.72	51.37	10.80	8.90
6:0	128	124	49.63	50.36	10.31	9.51
6:6	118	114	49.76	50.25	9.64	10.11
7:0	136	131	51.32	48.47	9.92	9.51
7:6	117	114	49.32	50.77	10.16	9.48

Table 29 Mean T scores of boys and girls at each age level.

5.00 First and second half-year testing

The results of one-way analyses of variance (Table 30) confirm the null hypothesis that there is no significant difference between the mean scaled BTBC scores (Form A) of children tested during the first half of their school year and those tested during the second half (as defined by September to the start of the February half-term break and from the end of the February break to the end of the summer term for all children on roll in September. For children admitted to nursery or reception classes in January, the second half of their school year is defined as starting on June 1 and for April admissions on November 1.)

Total sample	N	Mean T score	SD
1st half-year	945	49.89	9.87
2nd half-year	983	50.10	9.94

F = .22 df 1927; NS

Table 30 ANOVA: T scores of 1st and 2nd half-year testing.

6.00 Alternate form equivalence

The mean product moment correlation coefficient for the raw BTBC scores of 144 subjects (aged 4, 5, 6 and 7 years) each tested with Form A and Form B is .80. Table 32 shows the results for each age group and for the whole group. Means and standard deviations are in raw score units.

Age	r_{ab}	Mean A	Mean B	SD_A	SD_B	N
4:0	.81	22.94	24.06	8.14	6.95	36
5:0	.83	32.27	32.21	7.37	6.56	34
6:0	.78	39.76	39.34	4.99	4.76	38
7:0	.82	42.78	42.39	4.81	6.19	36
Mean	.80	34.56	34.65	10.16	9.41	144

Table 32 Correlation coefficients of Forms A and B.

6.01 Table 33 shows the proportion of children who score the same on each item on both forms together with the facility values for each form..

Item	Form A Facility	Form B Facility	Proportion with identical scores on each form
1	.97	.98	.95
2	.94	.93	.90
3	.82	.98	.83
4	.95	.96	.93
5	.96	.94	.92
6	.88	.66	.70
7	.97	.89	.93
8	.76	.78	.82
9	.87	.65	.78
10	.95	.67	.68
11	.92	.93	.94
12	.78	.71	.75
13	.85	.76	.91
14	.76	.65	.77
15	.70	.65	.81
16	.97	.97	.96
17	.65	.64	.87
18	.74	.85	.78
19	.75	.76	.78
20	.84	.88	.82
21	.81	.83	.90
22	.78	.71	.76
23	.54	.62	.74
24	.93	.21	.31
25	.55	.39	.65
26	.40	.43	.74
27	.80	.83	.85
28	.55	.98	.56
29	.60	.79	.67
30	.85	.82	.91
31	.51	.61	.71
32	.67	.69	.85
33	.49	.35	.84
34	.87	.65	.74
35	.58	.68	.61
36	.53	.94	.64
37	.47	.48	.77
38	.46	.56	.69
39	.56	.77	.74
40	.92	.94	.92
41	.76	.83	.78
42	.73	.70	.77
43	.48	.44	.72
44	.40	.52	.74
45	.32	.31	.84
46	.42	.42	.90
47	.44	.44	.75
48	.42	.44	.72
49	.40	.32	.82
50	.31	.37	.75

Table 33 Consistency of scoring on both Forms (A and B) for each item.

7.00 Construct validity

Principal components analyses, carried out on the correlation matrices of the BTBC item scores (Form A) with Varimax and Promax rotations, yielded no clear major factors. Table 34 shows the large numbers of First Order factors at each age level with the corresponding percentage variances. The first row shows the percentage variance accounted for by those principal components with eigen values greater than one at each age level. The second row shows the percentage variance after rotation.

7.01 Neither Varimax nor Promax rotations produced consistent factor loadings on the test items across the age levels. No meaningful interpretation could be made of the nature of the factors. Table 35 presents examples of factor loadings for the first 5 rotated components from a sample age group (5:0).

Factor	Age 3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6
1	12.6	11.7	15.2	15.4	13.2	11.1	12.6	9.0	14.0
	4.1	5.6	4.6	5.1	6.1	3.4	3.6	4.1	14.0
2	7.6	6.2	4.4	4.6	4.4	4.8	6.0	4.8	9.4
	7.8	5.7	5.2	5.5	4.2	5.0	6.0	4.6	4.0
3	5.0	4.0	3.8	3.6	3.8	4.2	5.4	4.3	4.4
	4.6	3.3	4.3	4.7	3.2	4.3	6.6	4.5	3.7
4	4.1	3.8	3.4	3.3	3.5	3.6	4.3	3.5	4.1
	4.2	3.4	2.8	3.0	3.3	3.0	4.4	3.2	4.0
5	3.9	3.5	3.4	3.2	3.3	3.5	3.7	3.5	4.0
	3.2	4.0	4.0	3.2	3.8	4.1	2.8	3.3	4.0
6	3.7	3.3	3.2	3.1	3.2	3.4	3.3	3.3	3.5
	3.4	3.3	3.2	4.0	3.3	3.4	3.2	3.3	3.3
7	3.5	3.3	2.9	3.0	3.1	3.3	3.2	3.1	3.3
	3.2	2.9	2.9	3.3	3.0	3.8	2.9	3.5	2.8
8	3.4	3.1	2.9	2.8	3.0	3.1	3.0	2.9	3.2
	3.6	2.7	3.3	2.9	2.6	3.3	3.3	2.9	3.4
9	3.3	3.0	2.8	2.7	2.7	3.0	2.9	2.8	3.0
	3.9	3.9	3.6	3.0	3.3	3.3	3.8	2.8	2.9
10	3.0	2.9	2.7	2.6	2.6	2.7	2.8	2.8	3.0
	3.4	3.0	4.7	2.8	3.8	3.4	3.3	2.7	3.0
11	2.8	2.7	2.6	2.6	2.6	2.7	2.7	2.6	2.9
	3.1	3.7	3.5	2.7	2.8	3.0	3.1	4.3	4.2
12	2.8	2.6	2.6	2.4	2.5	2.6	2.6	2.5	2.8
	4.1	3.0	3.3	3.7	3.0	2.8	3.5	2.7	2.8
13	2.7	2.6	2.5	2.3	2.3	2.5	2.5	2.4	2.5
	3.9	2.9	4.2	2.6	2.8	2.8	3.8	2.8	3.5
14	2.5	2.4	2.4	2.2	2.3	2.5	2.4	2.4	2.5
	2.8	3.4	3.2	2.7	3.7	2.8	2.6	2.7	3.3
15	2.3	2.3	2.3	2.2	2.2	2.4	2.3	2.4	2.5
	2.9	3.3	2.9	3.9	2.6	3.7	3.2	2.4	3.1
16	2.2	2.3	2.1	2.1	2.2	2.2	2.2	2.3	2.4
	3.3	2.8	2.8	3.1	3.6	3.0	3.0	3.0	3.2
17	2.1	2.2	2.1	2.1	2.1	2.1	2.2	2.2	2.4
	2.8	3.1	2.8	3.0	2.9	2.9	3.3	2.6	3.0
18	2.0	2.1	-	2.0	2.0	2.1	2.1	2.1	2.2
	3.6	3.4	-	3.2	3.3	2.9	3.9	2.9	2.9
19	2.0	2.0	-	-	2.0	2.0	-	2.1	2.0
	3.7	2.7	-	-	2.0	2.5	-	2.7	3.0
20	-	-	-	-	-	-	-	2.0	-
	-	-	-	-	-	-	-	2.0	-
21	-	-	-	-	-	-	-	2.0	-
	-	-	-	-	-	-	-	2.0	-
Total variance	71.4	66.0	61.3	62.3	63.1	63.5	66.1	65.0	74.0

Table 34 First Order factors: percentage variance for each factor at each age level.

Item	Factor 1 V*	Factor 1 P**	Factor 2 V	Factor 2 P	Factor 3 V	Factor 3 P	Factor 4 V	Factor 4 P	Factor 5 V	Factor 5 P
1			-.59	-.73						
2							.34	.38		
3					-.60	-.72				
4			-.45	-.42	-.26			.26		
5				-.27					.26	
6					-.30	-.25				
7			-.41	-.42						
8			-.28							-.31
9			-.61	-.71	-.25					
10										
11			-.72	-1.1						
12			-.34	-.45						
13								.28		
14	-.32				-.39	-.31				
15	-.40	-.34							-.28	-.29
16										
17	-.69	-.80								
18									-.37	-.30
19										
20			-.27	-.29	-.52	-.64				
21			-.34		-.37	-.29				
22			-.34	-.36						
23					-.64	-.87				
24										
25				.33						
26										
27										
28						-.25				
29									-.74	-.97
30	-.33						.32	.37	-.26	
31					-.50	-.58				
32	-.69	-.76								
33	-.47	-.44								
34										
35										-.26
36										
37										
38										
39										
40										
41	-.31	-.25			-.27					.29
42	-.35	-.23					-.49	-.51		
43	.32	-.49								
44							-.71	-.84		
45			-.28	-.33		-.28				
46										
47										
48									-.30	-.35
49	-.32	-.25								
50				.23						

* Varimax ** Promax

Table 35 Sample of factor loadings (age 5:0).

8.00 Content Validity

Teachers' discourse

During 11 hours 8 minutes of observed teacher discourse, involving eleven teachers in two schools (see Chapter 8, iii, 4.00 - 4.06), the following observations were made:

Thirteen of the original 50 BTBC expressions were never used with any age group between 3:0 and 8:0. These were:

some, not many
farthest
widest
several
not first or last
below
medium-sized*
right
zero**
separated
skip**
equal
least

* the alternative 'middle-sized' was not used either

** the words of the present adaptation viz. 'without any' were used and the past tense of 'miss' was used.

A further 5 BTBC expressions were only used with 7 year olds:

between
nearest
almost

centre

forward

'Above' and 'pair' were only used with 6 and 7 year olds.

A further 5 terms were only used in reception classes:

whole

alike

as many

in order

row

A further 6 BTBC expressions were not heard in either of the two nurseries:

away from

around

second

third

behind

always

'Beginning' and 'matches' were only heard in the nursery whilst 'left' was used with nursery children and the 6 to 7 year olds.

8.01 The remaining 15 BTBC terms were heard in the nurseries and in all or most of the other classes:

top

through

next to

inside

middle

few

over

most

corner

different

after

half

side

other

every

38% of the terms of the original BTBC items were observed in use with nursery children of 3 and 4 years

44% were observed in use with children aged 4 and 5

32% " " " " " " " 5 and 6

44% " " " " " " " 6 and 7

50% " " " " " " " 7 years

Thus observation of this sample of teachers' discourse suggests that understanding the entire content of the BTBC (1971) is not necessary for success in nursery and first school classes in England.

8.02 Table 36 shows the frequency of observation of each whole BTBC (1971) term, each part of a BTBC term and each term as adapted in the present investigation.

x denotes that the term (or part) occurred on fewer than 6 occasions.

xx denotes that the term occurred between 6 and 15 times.

xxx denotes that the term occurred more than 15 times.

Item	Age:	3/4	4/5	5/6	6/7	7 - 8
1	Top	xx	x	xx	xx	xx
2	Through	x	-	x	x	x
3	Away from	-	-	x	x	x
4	Next to	x	x	x	x	-
5	Inside	x	x	x	xx	xxx
6	Some, not many	-	-	-	-	-
	Some	xxx	xx	xx	xx	xxx
	Not many	-	-	x	-	-
7	Middle	x	x	-	-	x
8	Few	x	-	x	x	x
9	Farthest	-	-	-	-	-
10	Around	-	x	-	xx	x
11	Over	x	x	x	xx	x
12	Widest	-	-	-	-	-
13	Most	x	-	-	x	x
14	Between	-	-	-	-	x
15	Whole	-	x	-	-	-
16	Nearest	-	-	-	-	x
17	Second	-	x	x	-	x
18	Corner	x	-	-	x	x
19	Several	-	-	-	-	-
20	Behind	-	x	x	x	x
21	Row	-	x	-	-	-
22	Different	x	x	x	xx	xx
23	After	x	x	x	x	xx
24	Almost	-	-	-	-	x
25	Half	x	x	-	x	x

Table 36

continued over...

Item	Age: 3/4	4/5	5/6	6/7	7 - 8
26 Centre	-	-	-	-	x
27 As many	-	x	-	-	-
28 Side	xxx	x	-	x	xxx
29 Beginning	x	-	-	-	-
30 Other	xxx	xx	x	xx	x
31 Alike	-	x	-	-	-
32 Not first or last	-	-	-	-	-
First	xx	xx	xx	xx	xx
Last	-	x	x	x	xx
33 Never	x	x	x	x	x
34 Below	-	-	-	-	-
35 Matches	x	-	-	-	-
36 Always	-	x	x	x	x
37 Medium-sized/ middle-sized	-	-	-	-	-
38 Right	-	-	-	-	-
39 Forward	-	-	-	-	x
40 Without any	-	x	-	-	-
Zero	-	-	-	-	-
41 Above	-	-	-	x	-
42 Every	x	x	x	xx	xx
43 Separated	-	-	-	-	-
44 Left	x	-	-	x	-
45 Pair	-	-	-	x	-
46 Skip	-	-	-	-	-
Miss	-	-	-	-	-
47 Equal	-	-	-	-	-
48 In order	-	x	-	-	-
49 Third	-	x	x	-	-
50 Least	-	-	-	-	-

Table 36 Frequency of BTBC terms observed in teachers' discourse.

8.03 Table 37 shows the frequency of occurrence, in the observed sample, of synonyms, derivatives and parts of the BTBC terms used with the same meaning as that tested by the BTBC. (x is used as in table 36.)

Item	Related word(s)	3/4	4/5	5/6	6/7	7 - 8
3	Away	xx	x	x	x	x
	Put/clear away	x	xx	xx	xx	x
	Wash away	x	-	-	-	-
	Sent away	–	x	-	-	-
4	Next (place)	x	x	-	x	xx
	Close to/by	-	-	x	x	x
5	In	xxx	xxx	xxx	xxx	xxx
6	Some	xxx	xx	xx	xxx	xxx
	*Many	x	x	x	x	xx
9	Far, farther, further	x	x	x	x	x
	Distance	x	-	-	-	-
10	Round	x	x	xx	xx	xx
12	Wide, wider	x	-	-	x	xxx
13	More	xxx	xxx	xxx	xxx	xxx
15	All	xxx	xxx	xxx	xxx	xxx
	Completely	-	-	-	-	x
16	Near	-	-	-	x	-
	(see also 'close': item 4 above)					
20	At the back	x	-	x	x	x
24	Nearly	xx	xx	x	x	x
	Not quite	-	-	-	-	x
25	Semi-circle	x	-	-	-	-
	Semi-detatched	-	-	-	x	-
26	Middle	x	x	-	-	x
27	So many	-	-	x	-	-
	Too many	-	-	x	-	-
	How many?	xx	xxx	xxx	xx	xxx

Table 37 Continued over...

Item	Related word(s)	3/4	4/5	5/6	6/7	7 - 8
29	Begins, began, begun	x	x	-	x	-
	Starts, started	xx	xx	-	xx	xx
	Starting	-	-	-	-	x
30	Another	xxx	xxx	xxx	xxx	xxx
31	Like	xxx	xx	xx	xx	xxx
	Same, same as	xxx	x	-	xx	xx
	Similar	x	-	-	-	-
34	Under	xx	x	-	xx	xx
	Underneath	xx	x	x	-	x
	Beneath	-	-	-	-	x
35	Goes with	x	-	-	-	-
	Belongs	-	x	x	-	-
36	Incessant	-	-	-	-	x
37	Size	x	-	-	x	x
	Big, bigger, larger	xxx	xxx	xx	xx	xx
	Small, little	xxx	xxx	xx	xx	xx
	Smaller	-	-	x	-	x
39	Bending	-	-	-	-	xxx
40	*Without	-	x	-	-	x
	*Any	xx	xxx	xx	x	xxx
	*Not any	-	-	-	x	-
42	Each	x	x	-	x	xx
43	Separately	-	-	-	-	x
45	Both	xx	-	-	x	-
	Double	x	-	-	-	-
	Partner	-	-	-	xxx	-
	You two	-	-	-	-	x
46	Missing, missed out	x	-	x	x	x
47	(see item 31)					
50	Less	-	-	x	x	x

* these frequencies are not included in Table 36

Table 37 Frequency of synonyms, derivatives and parts of BTBC terms observed in teachers' discourse.

8.04 Table 38 shows the occurrence of antonyms of the BTBC terms in the observed sample. (x is used as in tables 36 and 37.)

Item	Antonym	3/4	4/5	5/6	6/7	7 -
1	Bottom	x	x	x	x	xx
5	Outside	x	x	x	xx	x
6	Single	x	-	-	-	-
	Total number	-	-	-	-	xx
8	Many	x	x	x	x	xx
	Lot(s)	x	-	-	-	-
	Plenty	x	-	x	-	-
12	Narrows	-	-	-	-	x
15	Part	-	-	-	x	x
	Piece	-	x	-	-	xx
	Compartment	x	-	-	-	-
20	In front	x	-	-	x	x
23	Before (temporal)	x	x	x	xx	xxx
29	Finished	xx	xx	xx	-	xx
	Completed	-	-	-	-	x
	End	x	-	x	-	x
	Stop	x	xxx	xxx	xxx	xxx
39	Backward(s)	x	-	-	-	x
	Back to	xx	xx	xx	x	x

Table 38 Frequency of antonyms of BTBC terms observed in teachers' discourse.

Any item which has another item as its antonym is not included in table 38 viz. items 2, 4, 9, 10, 11, 13, 16, 22, 25, 28, 31, 33, 34, 36, 38, 40, 41, 42, 44, 50.

8.05 BTBC terms which were used, in the observed sample, with a meaning different from that tested by the BTBC were:

after as in 'look after'

away as in 'right away'

right as in 'correct' and as in 'right away'

around (round) as in 'turn around', 'right/wrong way round', 'changed round'

never as in 'she/they never' (i.e. she didn't)

next as in 'next time', 'do ... next'

over as in 'fall over', 'over and over again', 'knock over', 'change over', 'turn over'

some as in the singular sense viz. somebody, someone, something

some as in 'somehow'

8.06 The evidence in Tables 36, 37 and 38 shows that, in the observed sample of teachers' discourse, only one of the terms used in the present adapted version of the BTBC is not represented in any form either in its BTBC form or by any related word or words.

Thus the observed sample of teachers' discourse suggests that most of the content of the present adapted version of the BTBC is relevant to nursery and first school classrooms in England. However, nomenclature different from that of the BTBC may be used by teachers when referring to a BTBC concept. Item 19 ('several') is likely to be irrelevant for most children in nurseries and first schools in England.

School Texts - Reading Books

8.07 In seven commonly used graded reading series[1] 6 of the 50 concept terms, as used in the present investigation, appear in each of the series;

a further 5 terms appear in 6 of the series

"	"	6	"	"	"	5	"	"	"
"	"	3	"	"	"	4	"	"	"
"	"	5	"	"	"	3	"	"	"
"	"	7	"	"	"	2	"	"	"
"	"	10	"	"	"	1	"	"	"

8 terms do not appear in any of the series, neither do the discarded words 'skip' and 'zero'. Two further words used in the BTBC, 'empty' (Form A) and 'full' (Form B), appear in five of the series.

If the supposition, that most schools do not confine their pupils to books from only one graded series, is correct then the above evidence suggests that 84% of the BTBC terms might be encountered by children whilst reading books from graded schemes. On the other hand, they are unlikely to meet 16% of the BTBC terms when reading such books.

8.08 The frequency of occurrence of BTBC terms, parts and derivatives, in each of the graded reading series, is shown in Table 39. In every case each word in context bears the same meaning as that tested by the BTBC.

[1] Details of the texts are in Appendix 13. See also Chapter 8, iii, 5.00.

		No. of books examined	No. of levels
A	Language in Action	19	4
B	Breakthrough to Literacy	24	4
C	Sparks	15	6
D	Dominoes	36	6
E	Ladybird	15	9
F	Reading 360	39	8
G	Through the Rainbow	36	10

Key to Table 39

x word(s) appears approximately 1 - 4 times at one or more levels.

xx word(s) appears approximately 5 - 10 times at one or more levels.

xxx word(s) appears over 10 times at one or more levels.

. appearing at 2 levels

.. appearing at 3 or more levels

... appearing at all or most levels (as defined by colour or number code in the handbook to the scheme.)

Item		Graded Series A	B	C	D	E	F	G
1	Top	xx ..	x ...	xx ..	xx ..	xx ..	xxx ..	xxx ..
2	Through		x .	x	xxx .	xx ..	xxx .	xxx ...
3	Away from	x			x	x ..	x	x ..
	Away	xx ...	x ...	xxx ..	xxx ..	xxx ..	xxx ..	xxx ...
4	Next to			x			x	
	Next (spatial)	x	x .	x	x .	xx ..	xxx ..	xxx ..
5	Inside	x .	x	xx .	xxx .	xx ..	xxx ..	xxx
6	Some, not many							
	Not many		x			x		x
	Some	x .	xxx ...	xxx ..	xxx ...	xxx ...	xxx ..	xxx ...
	Something etc	x .	xx ...	xx ..	xxx ..			x .
7	Middle				x		x .	xx .
8	Few	x		x	x		x	
9	Farthest							
	Far, farther	x .		x	xxx	x	xx .	xxx ...
10	Around	x		xx	xxx .	x	xxx ..	x
	Round		xx .	xx	xxx ..	xxx ..	xxx ..	xxx ...
11	Over	xx ...	x ...	xx .	xxx ...	xxx ..	xxx ..	xxx ..
12	Widest							
	Wide, wider		x		x		x ..	x
13	Most				x	xx ..	x	
14	Between			x	x	xx .	x	x
15	Whole						x .	
16	Nearest						x	
	Near, nearer	x .	x .	xx	xx ..	x ..	x	xx
17	Second	x			x .		xx	x ...
18	Corner			x			xx .	
19	Several				x			x
20	Behind	x ...	x	x	xx .		xxx .	xxx .
21	Row						x	
22	Different	xx	x					x
23	After	x .	x ...	xx ...	xx ..	xxx ..	x .	xxx ...
24	Almost						x .	

Table 39

continued over...

		A	B	C	D	E	F	G
	Empty		x	x	x	x		xx ..
	Full	x		x .	x	x	x	x
25	Half						x	xx .
26	Centre						x	
27	As many					x		
	Many	x	x	x ..		xxx ..	x ..	xx .
28	Side	x .	x .	x	x .		x .	xxx ..
29	Beginning	x	x				x	
	Begin/s/an/un			x	xxx .			xxx .
30	Other(s)	x	x .	xx .	xxx ..	xxx ..	x	xxx .
31	Alike	xx				xx		
	Like	x ...	x	xx .	xxx ..	xx ..		
32	Not first or last					x		
	First	x	x .	xx	xx ..	xxx ...	xx .	xxx ...
	Last	x .	x	x .	x .	x	x .	xxx .
33	Never	x ...	x		xxx	xx ..	x	xxx .
34	Below	x					x	
35	Matches							
	Match	x						
36	Always		xx .	x	xx ..	xx .	x	
37	Medium/middle sized							
38	Right	x	xx .		x		x	xx
39	Forward							x
40	Without any					x	xx	x
	Without			x	x .		x	x .
	Any	x	x	xx	xxx	xx ..	x ..	xxx .
	Anything etc		x				xx .	
41	Above	x .					x .	xx ..
42	Every	x	x .	xx	xxx ...	xx ..	x ..	xxx ..
	Everything etc	x .	x .	x .	xxx ..	x .	xx .	xx ..
43	Separated							
44	Left	x .	xx		x		x	xx
45	Pair		x					x
46	Miss							xx

Table 39

Continued over...

		A	B	C	D	E	F	G
	Missing, missed							x
47	Equal							
48	In order							
49	Third	x		x	x .		x	x ...
50	Least						x	

Table 39 Occurrence of BTBC terms, parts and derivatives in reading books.

Table 39 shows only three of the 50 concept terms as absent, both in whole and in part, from the books examined viz. 'equal', 'in order' and 'separated'. The absence of 'skip' and 'zero' supports the abandonment of these words. A further 16 terms, both in whole and in part, are seldom represented.

The frequency with which the remaining 31 BTBC terms occur, either in whole or in part, suggests that 62% of the BTBC terms are relevant to the content of graded reading series used in first schools.

8.09 Examples of the BTBC terms used in graded reading books where the meaning differs from that tested by the BTBC are:

after as in 'look after'

around as in 'turned around', 'around the house' (i.e. in various parts of the house)

away as in 'straight away', 'grinning away' left as in '... sandwiches left', 'left it'

right as in 'hold it right', 'right place', 'right up me', 'right on time'

next as in 'next Saturday'

over as in 'the party is over', 'play is over'

'I like' appears frequently (Cf item 31 'alike').

School Texts - Mathematics Books

8.10 None of five sets of pupils' mathematics workbooks[1] contains the following 7 Boehm concept terms:

some, not many

farthest

almost

centre

alike

not first or last

medium/middle sized

A further 7 terms appear in one or two sets but with only 5 or fewer occurrences:

away from

several

never

below

always

above

separated

[1] See Chapter 8, iii, 5.01 and Appendix 14.

left as in '... sandwiches left', 'left it'

right as in 'hold it right', 'right place', 'right up me', 'right on time'

next as in 'next Saturday'

over as in 'the party is over', 'play is over'

'I like' appears frequently (cf item 31 'alike').

School Texts - Mathematics Books

8.10 None of five sets of pupils' mathematics workbooks[1] contains the following 7 Boehm concept terms:

some, not many

farthest

almost

centre

alike

not first or last

medium/middle sized

A further 7 terms appear in one or two sets but with only 5 or fewer occurrences:

away from

several

never

below

always

above

separated

[1] See Chapter 8, iii, 5.01 and Appendix 14.

A further 8 terms also appear in only one or two of the sets but with more frequent occurrence:

middle

whole

nearest

behind

beginning

forward

pair

widest

The remaining 28 BTBC items occur frequently in all or most of the sets of workbooks.

8.11 Only two of the BTBC terms ('several' and 'separated') do not appear either in whole or in part, in the mathematics vocabulary recommended by the mathematical handbooks for teachers.[1] As both these terms occur, though rarely, in the workbooks examined, it may be assumed that the content of the BTBC (as adapted in the present investigation) is relevant to mathematics in first schools.

8.12 Furthermore, only three BTBC terms (as presently adapted) are absent, in whole or in part, from the mathematical vocabulary suggested to nursery teachers. These three terms are 'alike', 'left' and 'equal'. Therefore, 94% of the content of the BTBC (as presently adapted) appears to be relevant to nursery education.

[1] See footnote to previous page.

School Science

8.13 The early stages of the project "Science 5/13" contain all but four of the BTBC terms, in whole or in part, (as used in the present adaptation) : these are 'whole', 'separated', 'miss' and 'third'. Thus at least 92% of the content of the BTBC (as used in the present adaptation) appears to be relevant to science education in first schools.

Nursery Assessment

8.14 The oral directions and questions used in "Assessment in Nursery Education" (Bate and Smith, 1978) contain 20 of the BTBC terms exactly as used in the present adaptation and 6 further terms in part.

24 of the BTBC terms do not feature though some of the concepts are used but with different vocabulary viz.

'line' instead of 'row'

'like' and 'same' instead of 'alike'

'belong(s) instead of 'matches'

These observations suggest that at least half the concepts of the BTBC (as used in the present adaptation) might be "commonly found among children" in nurseries (Bate and Smith, 1978, p.9).

8.15 Limited samples of observations of the speech of nursery children and teachers (see Chapter 8, iii, 6.00) in socially disadvantaged areas, yielded the results shown in Table 40 which shows the occurrence of BTBC terms, or their parts, and the usage of other words where these were used more frequently than the BTBC vocabulary.

Item		Teachers' speech		Children's speech	
		frequency	other word(s)	frequency	other word(s)
1	Top	2			
3	Away from				
	Away			3	
4	Next to	1			
	Next (spatial)	1			
5	Inside	1	in		in
6	Some, not many				
	Some			17	'somat'
	Many	1		1	
10	Around		round		
11	Over	2		2	
15	Whole	1			
18	Corner			1	
20	Behind	2			
22	Different	2			
24	Almost		nearly		nearly
25	Half	2			
28	Side	7			
30	Other	6		1	
31	Alike		like, same		like
32	Not first or last				
	First	2		3	
	Last	2			
33	Never	4			
34	Below		under		under
40	Without any				
	Any	7		3	
41	Above	1			
42	Every	3		1	
44	Left	1			

Table 40 Sample of speech in nurseries (Smith, [illegible]).

The 27 BTBC terms not included in Table 40 were not observed in this limited sample of speech.

Summary

8.16 Table 41 shows the occurrence of BTBC terms (as presently adapted) or their parts in each of the domains observed.

Item	Teachers	Children[1]	Reading	Maths	Science	Nursery Assessment
1 Top	x	-	x	x	x	x
2 Through	x	-	x	x	x	x
3 Away from	x	-	x	x	x	-
4 Next to	x	-	x	x	x	-
5 Inside	x	-	x	x	x	x
6 Some, not many	-	-	-	-	-	-
Some	x	x	x	x	x	x
Many	x	x	x	x	x	x
7 Middle	x	-	x	x	x	x
8 Few	x	-	x	x	x	-
9 Farthest	-	-	-	x	x	-
10 Around	x	-	x	x	x	-
11 Over	x	x	x	x	x	x
12 Widest	-	-	-	x	-	-
13 Most	x	-	x	x	x	-
14 Between	x	-	x	x	x	x
15 Whole	x	-	x	x	-	x
16 Nearest	x	-	x	x	-	x
17 Second	x	-	x	x	x	x
18 Corner	x	x	x	x	x	-
19 Several	-	-	x	x	x	-
20 Behind	x	-	x	x	x	x
21 Row	x	-	x	x	x	-
22 Different	x	-	x	x	x	x
23 After	x	-	x	x	x	x
24 Almost	x	-	x	x	x	-
25 Half	x	-	x	x	x	x
26 Centre	x	-	x	x	x	-
27 As many	x	-	x	x	x	-
28 Side	x	-	x	x	x	x

Table 41

Continued over...

[1] As explained earlier the sample of children's speech is very limited.

Item		Teachers	Children	Reading	Maths	Science	Nursery Assessment
29	Beginning	x	-	x	x	x	-
30	Other	x	-	x	x	x	x
31	Alike	x	-	x	x	x	-
32	Not first or last	-	-	x	-	-	-
	First	x	x	x	x	x	x
	Last	x	-	x	x	-	x
33	Never	x	-	x	x	x	-
34	Below	-	-	x	x	x	-
35	Matches	x	-	-	x	x	-
36	Always	x	-	x	x	x	-
37	Medium/middle sized	-	-	-	-	-	-
	Medium (weight)	-	-	-	-	x	-
	Middle	x	-	x	x	x	x
38	Right	-	-	x	x	x	x
39	Forward	x	-	x	x	x	-
40	Without any	x	-	x	x	-	-
41	Above	x	-	x	x	x	-
42	Every	x	x	x	x	x	-
43	Separated	-	-	-	-	-	-
	Separately	x	-	-	x	-	-
44	Left	x	-	x	x	x	x
45	Pair	x	-	x	x	x	-
46	Miss	-	-	x	x	-	-
47	Equal	-	-	-	x	x	-
48	In order	x	-	-	x	x	x
49	Third	x	-	x	x	-	-
50	Least	-	-	x	x	x	x

Table 41 Occurrence of BTBC terms in sampled curricula.

Only two of the BTBC terms (as presently adapted) are absent from all of the domains observed. 'Separated' was observed in none of the domains but 'separate' and 'separately' were observed.

'Medium/middle-sized' is absent but 'medium' and 'middle' appear. Moreover, the term 'middle-sized' is a significant feature of early childhood education (Parry and Archer, 1975).

Therefore, it can be concluded that the content of the BTBC (as presently adapted) can be considered relevant to first schools in England though some children may not be familiar with all the terms used.

9.00 Predictive validity

A product moment correlation matrix (with missing data) of the following variables is reproduced in Table 42a:

Variable 1. Sex (boys coded 1; girls 2).

Variable 2. Spoken language assessed by teachers approximately one year after the BTBC (Form A) (below average coded 1; average 2; above average 3).

Variable 3. Mathematics A (counting and computation) assessed and coded as for Variable 2.

Variable 4. Mathematics B (concepts) assessed and coded as above.

Variable 5. Reading assessed and coded as above. (See Chapter 8, iii, 3.00.)

Variable 6. BTBC T scores (unsmoothed).

Sex	Speech	Maths A	Maths B	Reading	BTBC T score
1	2	3	4	5	6
1	.065	.024	-.018	.279*	.029
2		.611*	.615*	.606*	.449*
3			.868*	.699*	.481*
4				.669*	.457*
5					.424*

Table 42a Correlation matrix of T scores, sex and teachers' assessments.

* highly significant ($p < \cdot 01$)

The number of observations, mean and SD for each variable is shown in Table 42b.

Variable	N	Mean	SD
1. Sex	1928	1.50	.50
2. Spoken language	428	2.15	.62
3. Maths A	401	2.04	.63
4. Maths B	371	2.04	.67
5. Reading	388	2.11	.67
6. T scores	1928	50.00	9.90

Table 42b Mean, N and SD of variables in Table 41B.

Social Class Differences

10.00 H_0 There are no significant differences between the BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations), group 2 (as defined by parents in other 'white collar' and skilled manual occupations) and group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment.

Table 43 shows the results of one-way analyses of variance of the T scores of those children who were known for certain to belong to SES groups 1, 2 and 3.

	Total sample	SES group 1	SES group 2	SES group 3
N	1498	325	721	452
Mean score	49.98	56.49	50.07	45.16
SD	28.98	9.29	9.16	9.23
F	14.70			
df	2/1495			
p <	.01	1 with 2	1 with 3	2 with 3
t		3.35	5.43	2.85
df		1044	775	1171
p <		.001	.001	.01

Table 43 ANOVA: T scores of SES groups 1, 2 and 3

As the F value is highly significant the null hypothesis is rejected.

10.01 As the t values are all highly significant the following hypotheses are accepted:

H_1 The BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations) are significantly higher than those of children in socio-economic group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment).

H_2 The BTBC mean scaled scores of children in socio-economic group 1 (as defined by parents in professional and managerial occupations) are significantly higher than those of children in socio-economic group 2 (as defined by parents in other 'white collar' and skilled manual occupations).

H_3 The BTBC mean scaled scores of children in socio-economic group 2 (as defined by parents in other 'white collar' and skilled manual occupations) are significantly higher than those of children in socio-economic group 3 (as defined by parents in semi- or unskilled occupations or in long-term unemployment).

10.02 Though the differences between the mean scaled scores of SES groups 1, 2 and 3 in the total sample are highly significant, t tests at each age level reveal significant differences only at ages 3:6, 5:0 and 5:6. Table 44 shows ANOVA results at each age level.

	3:6	4:0	4:6	5:0	5:6	6:0	6:6	7:0	7:6
Grand mean	50.65	49.46	50.49	50.04	49.81	49.53	49.87	50.50	49.74
SES 1 mean	59.64	55.50	59.74	55.64	57.75	55.24	55.50	56.83	55.06
SD	6.79	9.31	7.11	9.33	9.00	10.36	9.58	10.22	9.07
N	22	28	23	45	40	38	46	47	36
SES 2 mean	50.91	49.76	49.29	50.47	49.94	49.59	50.14	50.63	50.23
SD	9.19	8.75	9.15	9.53	8.38	9.56	10.24	8.03	9.55
N	35	54	76	95	86	106	87	93	89
SES 3 mean	44.74	44.95	47.39	45.45	44.05	45.54	44.44	45.25	44.96
SD	8.43	10.10	9.56	9.66	8.80	9.65	7.87	9.63	9.42
N	35	41	39	64	57	56	52	59	49
F	2.26	0.91	1.12	2.19	2.98	1.32	1.95	1.91	1.14
df	2/89	2/120	2/135	2/201	2/180	2/197	2/182	2/196	2/171
1 with 2 t	1.24	0.77	1.33	1.13	1.50	1.05	1.06	1.14	0.80
df	55	80	97	138	124	142	131	138	123
1 with 3 t	2.13*	1.35	1.42	2.08*	2.44**	1.63	1.97	1.95	1.50
df	55	67	60	107	98	92	96	104	83
2 with 3 t	1.00	0.73	0.29	1.13	1.27	0.86	1.17	1.07	0.96
df	68	93	113	157	141	160	137	150	136

* $p < .05$ ** $p < .02$

Table 44 ANOVA: T scores of SES groups 1, 2 and 3 at each age level.

10.03 Figure 15 depicts the mean scaled scores of each SES group by age level.

Figure 15 Mean scaled scores of each SES group by age level.

10.04 The ANOVA results do not justify acceptance of the hypothesis (H_4) that social class differences (as defined by parental occupations), which are reflected in the BTBC mean scaled scores, are greater below 5 years than at 7 or 7½ years of age.

10.03 Figure 15 depicts the mean scaled scores of each SES group by age level.

Figure 15 Mean scaled scores of each SES group by age level.

10.04 The ANOVA results do not justify acceptance of the hypothesis (H_4) that social class differences (as defined by parental occupations), which are reflected in the BTBC mean scaled scores, are greater below 5 years than at 7 or 7½ years of age.

Differences Between Domiciliary Areas

11.00 H_0 There are no significant differences between the BTBC mean scaled scores of children attending rural, inner city or 'other urban' schools (as defined by the DES, 1978).

Table 45 shows the results of one-way analyses of variance of the T scores between subjects attending schools in rural, urban and inner city areas.

	Total sample		Rural	Urban	Inner city
N	1928		727	961	240
Mean score	49.95		50.81	49.69	48.38
SD	9.90		9.78	9.97	9.77
			Rural with Urban	Rural with Inner City	Urban with Inner City
F	6.10	t:	2.31	3.30	1.84
df	2/1925		1686	965	1199
p <	.01		.05	.001	NS

Table 45 ANOVA: T scores of rural, urban and inner city children

11.01 As the F value is highly significant the null hypothesis is rejected.

11.02 Though t tests show the rural children to have achieved significantly higher T scores than both the urban and inner city children this is likely to be largely dependent on the socioeconomic composition of the three groups.

11.03 Therefore children from similar socioeconomic backgrounds (SES groups 2 and 3) were compared: Table 46 gives the results of one-way analyses of variance between the T scores of children attending schools in areas with marked social difficulties (rural and urban) and children in other rural and urban areas, including inner city areas.

	SES 2 and 3 Total Sample	SES 2 and 3 Areas of Disadvantage	SES 2 and 3 Other Areas
N	1312	302	1010
Mean score	48.13	46.15	48.72
SD	9.47	9.74	9.32
F	17.38		
df	1/1310		
p <	.001		

Table 46 ANOVA: T scores from disadvantaged and other areas.

11.04 As the F value in Table 46 is highly significant the following hypothesis is accepted:

H_5 The BTBC mean scaled scores of children in socio-economic groups 2 and 3 (as defined by parental occupations) attending rural or urban schools in areas with marked social difficulties (as defined by low incomes, high unemployment and lack of amenities) are lower than those of children in socio-economic groups 2 and 3 attending schools in areas without a preponderance of such difficulties.

11.05 Table 47 gives the results of one-way analyses of variance between the T scores of inner city children, children in urban rehousing estates, children in other urban areas, children in disadvantaged rural areas and children in other rural areas. Only those children known to belong to SES groups 2 and 3 are included.

SES 2 and 3	N	Mean SS	SD	
Total sample	1312	48.13	9.47	
1. Rural 'other'	305	49.89	9.10	F = 7.08
2. Urban 'other'	489	48.49	9.33	df = 1311
3. Inner city	216	47.59	9.44	$p < .01$
4. Rural disadvantaged	166	46.95	9.81	
5. Urban disadvantaged	136	45.16	9.59	

Table 47 ANOVA: T scores of SES groups 2 and 3 in five types of locations.

Table 48 shows the results of testing the significance of the differences between the T scores of the five locational groups.

Locational groups	t value	df	p
Group 1 with Group 2	2.03	792	.05
Group 3	2.75	519	.01
Group 4	3.24	469	.01
Group 5	4.88	439	.001
Group 2 with Group 3	1.18	703	NS
Group 4	1.83	653	NS
Group 5	3.66	623	.001
Group 3 with Group 4	.66	380	NS
Group 5	2.36	350	.02
Group 4 with Group 5	1.65	300	NS

Table 48 Differences between mean T scores of five locational groups.

The following hypothesis is thus accepted:

H_6 The BTBC mean scaled score of children attending inner city schools (as defined by the DES, 1978) is significantly higher than that of children attending schools in urban rehousing estates (as defined by post-war slum clearance).

11.07 Furthermore, the following additional hypotheses are accepted.

H_{13} The BTBC mean scaled scores of children attending inner city and other urban schools (as defined the the DES, 1978) and of children attending schools in rural areas with marked social difficulties (as defined by low incomes, high unemployment and lack of amenities) are significantly lower than the mean scaled scores of children attending other rural schools (as defined by the DES, 1978).

H_{14} The BTBC mean scaled scores of children from non-professional and non-managerial classes (as defined by parental occupations) who attend school in urban rehousing estates (as defined by post-war slum clearance), outside the city centre, is significantly lower than that of children from the same social class who attend schools in other urban areas awav from a city centre (as defined by the DES, 1978).

12.00 H_0 There is no significant difference between the number of correct responses to a BTBC term which is unmarked (positive) and the number of correct responses to the opposite marked (negative) term, as defined by Clark (1977).

Table 49 shows the responses to five pairs of Form A items where each pair consists of an unmarked term and its marked opposite. (Number of subjects = 1928)

Table 50 shows the equivalent results for Form B. (Number of subjects = 144)

FORM A

Unmarked term	N. right	P	Marked term	N. right	P
4. next to	1822	.95	3. away from	1567	.81
13. most	1674	.87	50. least	692	.36
15. whole	1445	.75	25. half	1211	.63
16. nearest	1892	.98	9. farthest	1648	.86
38. right	1088	.56	44. left	940	.49
Mean	1584.2	.82	Mean	1211.6	.63

t (correlated means) = 26·92; $p < .001$

Table 49 Responses to unmarked/marked pairs on Form A.

FORM B

Unmarked term	N. right	P	Marked term	N. right	P
4. next to	137	.95	3. away from	141	.98
13. most	125	.87	50. least	85	.59
15. whole	93	.65	25. half	68	.47
16. nearest	142	.99	9. farthest	95	.66
38. right	88	.61	44. left	98	.68
Mean	117	.81	Mean	97.4	.68

t (correlated means) = 10.05; $p < .001$

Table 50 Responses to unmarked/marked pairs on Form B.

As the mean differences between the responses to the unmarked and marked terms are highly significant the null hypothesis is rejected.

12.01 H_7 The number of correct responses to a BTBC term which is unmarked (positive) is significantly higher than the number of correct responses to the opposite term which is marked (negative), as defined by Clark, 1977.

H_7 is accepted for both Form A and Form B.

12.02 H_0 There is no significant difference between the frequency of (a) the larger object (or greater amount) chosen in error and (b) other errors on each BTBC item where this type of error is possible.

Tables 51 and 52 show the number of times each wrong response was recorded[1] for the total age-range together with the chi-square values calculated using the chi-square one-sample test. The wrong pictures or objects are numbered left to right or top to bottom unless otherwise indicated. The larger object or amount is underlined. Table 51 shows the errors[1] on Form A (N = 1928) and Table 52 shows the errors[1] on Form B (N = 144).

[1] Not every response is included in these tables. Some children made no choice by saying "Don't know", by pointing to more than one picture or, with the youngest children and the later, more difficult items, because the examiner omitted these items after successive and excessive failures. Some wrong responses were omitted because of ambiguous recording.

Form A Item	N. Errors pictures 1	2	3	χ^2	df	P
6.	200	61		74.03	1	.001
8.		226	188	3.49	1	NS
See also Table 53						
9.	55		207	88.18	1	.001
24.	94	27		37.10	1	.001
25.		390	332	4.66	1	.05
37.		554	239	125.13	1	.001
40.		73	63	0.74	1	NS
45.	1169	85		937.05	1	.001
50.		244	724	238.02	1	.001
Picture	1	3	4			
17.	225	97	157	51.36	2	.001
See also Table 53						

Table 51 Errors on Form A items where the larger amount may dominate.

Form B Item	N. Errors pictures 1	2	3	χ^2	df	P
6.		5	45	32.00	1	.001
8.		10	16	1.38	1	NS
See also Table 53						
21.		12	13			NS
24.		107	2	101.15	1	.001
37.		61	17	24.82	1	.001
40.		3	4			NS
45.	70	8		49.28	1	.001
50.		17	42	10.59	1	.01

Table 52 Errors on Form B items where the larger amount may dominate.

12.03 Except for items 8 and 17 the errors show similar patterns at each age-level. Different patterns are seen with items 8 and 17 where different errors predominate as the children mature. Details are shown in Table 53.

	N. Errors					
Item 8	pictures			χ^2	df	P
Form A	1	2	3			
ages 3, 4 & 5		168	71	39.37	1	.001
ages 6 & 7		58	117	19.89	1	.001
Form B						
ages 3, 4 & 5		6	6			NS
ages 6 & 7		4	10	2.57	1	NS
Item 17	pictures					
Form A	1	3	4			
ages 3, 4 & 5	208	65	137	74.80	2	.001
ages 6 & 7	17	32	20	5.48	2	NS

Table 53 Errors on items 8 and 17 by age level.

12.04 The null hypothesis is rejected for 9 of the 10 items on Form A (Table 51 and Table 53) and is accepted for item 40 where the difference between the number of error choices is non-significant.

The null hypothesis is rejected for 5 of the 8 items on Form B (Table 52 and Table 53) and is accepted for items 8, 21 and 40 where the difference between the number of error choices is non-significant.

12.05 H_8 The frequency of (a) the larger object (or greater amount) chosen in error will be higher than (b) other errors on each BTBC item where this type of error is possible.

Inspection of Tables 51 and 53 reveals that the number of times the greater amount is chosen in error is significantly higher on Form A items 6, 9, 17, 25, 37, 45, 50 and, for children under 6 years, item 8. On item 24A the frequency of the greater amount error is significantly lower than the other type of error.

Inspection of Tables 52 and 53 reveals that the number of times the greater amount is chosen in error is significantly higher on Form B items 24, 37, 45 and 50. On item 6B the frequency of the greater amount error is significantly lower than the other type of error.

Therefore H_8 is accepted for eight Form A items and four Form B items. H_8 is rejected for one Form A item and one Form B item.

12.06 H_0 On each item 34 and 41 (Form A) there is no significant difference between the wrong responses chosen.

Table 54 shows the number of times each wrong response was recorded for the total age-range together with the chi-square values calculated using the chi-square one-sample test. (See footnote on p. 277.)

Form A item		N. errors		χ^2	df	P
	on	over	edge			
34. below	208	13	4	354.32	2	.001
		below	side			
41. above		245	195	5.68	1	.02

<u>Table 54</u> <u>Errors on Form A items 34 and 41</u>.

The null hypothesis is rejected.

12.07 H_9 On each item 34 and 41 (Form A) the frequency of wrong responses on the vertical axis will be higher than the frequency of wrong responses chosen from the horizontal axis.

H_9 is accepted.

12.08 H_0 On items 26 and 29 there is no significant difference between the wrong responses chosen.

Item		N. Errors			χ^2	df	P
		top	other 1	other 2			
26	Form A	428	121	272	172.19	2	.001
26	Form B	37	7	35	21.39	2	.001
29	Form A	194		672	263.84	1	.001
29	Form B	15	13				NS

Table 55 Errors on items 26 and 29.

There is no significant difference between the errors on item 29 Form B. Therefore the null hypothesis is accepted for item 29 Form B and rejected for item 29 Form A and item 36 on both Forms.

The chi-square value for item 26 Form B is highly significant but the difference between the number of times the topmost marble was chosen and the number of times the marble outside the circle was chosen is very small.

12.09 H_{10} On Form A items 26 and 29 the frequency of the topmost object as a wrong response is higher than any other error frequency.

Inspection of Table 55 reveals that H_{10} can be accepted for Form A item 26 only. H_{10} is rejected for Form A item 29.

12.10 H_0 On item 9 there is no significant difference between the wrong responses chosen.

Item 9	N. Errors		χ^2	df	P
	nearest	other			
Form A	207	55	88.18	1	.001
Form B	38	11	14.88	1	.001

Table 56 Errors on item 9.

The null hypothesis is rejected.

12.11 H_{11} On item 9 the frequency of the nearest object as a wrong response is higher than any other error frequency.

Inspection of Table 56 reveals highly significant differences between the number of times the nearest object is chosen as a wrong response on both Form A and Form B. Therefore H_{11} is accepted.

12.12 H_0 The number of times the first object or picture in a row is chosen as an alternative wrong response to a BTBC item is significantly higher than the number of times any other response is chosen.

Tables 57 and 58 show the distribution of recorded errors on the items where the correct answer is not the first picture or object in a horizontal array. (See footnote on p. 277) Table 57 Form A is based on the recorded errors of 1928 subjects. Table 58 Form B is based on the recorded errors of 144 subjects.

Form A			N. Errors				χ^2	df	P
Items			pictures						
	1	2	3	4	5	6			
1.	60	24					15.43	1	.001
6.	200	61					74.03	1	.001
7.	74	64		31	40		23.21	3	.001
13.	113	121							NS
14.	33	130	122		86	88	72.31	4	.001
15.	303		134				65.36	1	.001
17.	225		97	157			51.36	2	.001
19.	219	56					96.62	1	.001
21.	131		139						NS
22.	263		213				5.25	1	1.05
24.	94	27					37.10	1	.001
30.	198		237				3.50	1	NS
31.	271		243	172	wrong pairs 155		44.12	3	.001
32.	229		366				31.55	1	.001
35.	468		222				87.70	1	.001
36.	255	494					76.26	1	.001
38.	575	128					284.22	1	.001
39	30	636					551.41	1	.001
42.	193		185						NS
43.	385	284					15.25	1	.001
45.	1169	85					937.05	1	.001
49.	321	370		152			93.10	2	.001

Table 57 First-in-line errors on 22 Form A items.

Table 57 reveals that four Form A items show no significant difference between the frequency of first-in-line errors and the frequency of other errors. Therefore the null hypothesis is accepted for Form A items 13, 21, 30 and 42. It is rejected for Form A items, 1, 6, 7, 14, 15, 17, 19, 22, 24, 31, 32, 35, 36, 38, 39, 43, 45, 49.

Form B	N. Errors					χ^2	df	P
Items	pictures							
	1	2	3	4	5			
1.	3		2					NS
5.	4	2						NS
7.	5	5		2	1	3.92	3	NS
10.	8	38				19.57	1	.001
12.	14	22				1.78	1	NS
13.	11		8					NS
14.	18	26				1.46	1	NS
17.	37	5		7		39.43	2	.001
22.	22		17					NS
23.	16	36				7.69	1	.01
30.	11		12					NS
				wrong pairs				
31.	13	13		10	11	2.92	3	NS
32.	24		18					NS
35.	20	19						NS
36.	4		2					NS
38.	50	2	4			28.14	2	.001
39.	19	9				3.57	1	NS
42.	12		24			4.00	1	.05
43.	44		18			10.90	1	.001
45.	70	8				49.28	1	.001
48.	16		29			3.76	1	NS
49.	30	28		9		19.19	2	.001

Table 58 First-in-line errors on 22 Form B items.

Table 58 reveals that 14 Form B items show no significant difference between the frequency of first-in-line errors and the frequency of other errors. Therefore the null hypothesis is accepted for Form B items 1, 5, 7, 12, 13, 14, 22, 30, 31, 32, 35, 36, 39, 48. It is rejected for Form B items 10, 17, 23, 38, 42, 43, 45, 49.

12.13 The number of times the first object or picture in a row is chosen as an alternative wrong response to a BTBC item is significantly higher than the number of times any other response is chosen.

Though errors on eighteen Form A items show significant differences, on four of these items (14, 32, 36, 39) the frequency of first-in-line errors is significantly lower than the frequency of other wrong responses. On item 49, though the frequency of the first-in-line error is significantly higher than the last-in-line error, the second child is chosen more often than the first child. Therefore H_{12} is rejected for Form A items 14, 32, 36, 39 and 49. H_{12} is accepted for the remaining thirteen items: items 1, 6, 7, 15, 17, 19, 22, 24, 31, 35, 43, 45.

Though errors on eight Form B items show significant differences, on items 10, 23 and 42 the frequency of first-in-line errors is significantly lower than the frequency of other wrong responses. Therefore H_{12} is rejected for Form B items 10, 23 and 42. H_{12} is accepted for Form B items 17, 38, 43, 45 and 49.

12.14 Tables 59 and 60 show the recorded errors, on Form A and Form B items which do not appear in Tables 51-58, together with chi-square values. (See footnote on p.277.)

Form A (1928 subjects)	N. Errors			χ^2	df	P
Items		pictures				
	1	2	3			
2.		59	23	15.81	1	.001
	on	under				
3.	75	245		90.31	1	.001
4.		51	17	17.00	1	.001
5.		51	51			NS
10.		66	74			NS
11.		45	47			NS
12.		249	123	42.68	1	.001
16	12	14				NS
18.		166	234	11.56	1	.001
		doll	boat			
20.		305	36	212.20	1	.001
23.		644	178	264.18	1	.001
27.	71	79 (The box 90)				NS
28.	601	242		152.88	1	.001
33.		327	362	1.78	1	NS
44.		102	651	400.27	1	.001
46.	853		20	794.83	1	.001
47.	126	276	39	195.55	2	.001
above are single responses. N. wrong pairs = 204.						
48.	392		306	10.60	1	.01

Table 59 Errors on 18 Form A items.

Form B (144 subjects)	N. Errors			χ^2	df	P
Items	pictures					
	1	2	3			
2.		5	4			NS
3.		1	2			NS
4.	rabbit 4	1	2			NS
11.	8		1	5.44	1	.02
15.		27	24			NS
16.	1	1				NS
18.		5	12	2.88	1	NS
19.		9	20	4.17	1	.05
20.	9		3	3.00	1	NS
25.		30	46	3.37	1	NS
27.	3	7	(The box 5)			NS
28.	2		2			NS
33.		36	26	1.61	1	NS
34.	26		23			NS
41.	11		15			
44.	6		40	25.13	1	.001
46.	50		5	18.41	1	.001
47.	4		21	11.56	1	.001

above are single responses. N wrong pairs = 10.

Table 60 Errors on 18 Form B items.

12.15 The predominating errors accepted by hypotheses 8 - 12 are depicted below. Figure 16a refers to Form A items and Figure 16b refers to Form B items.

Item numbers added in parentheses refer to items where the errors are in the direction predicted by the hypothesis but the errors are relatively few and do not reach statistical significance.

Form A

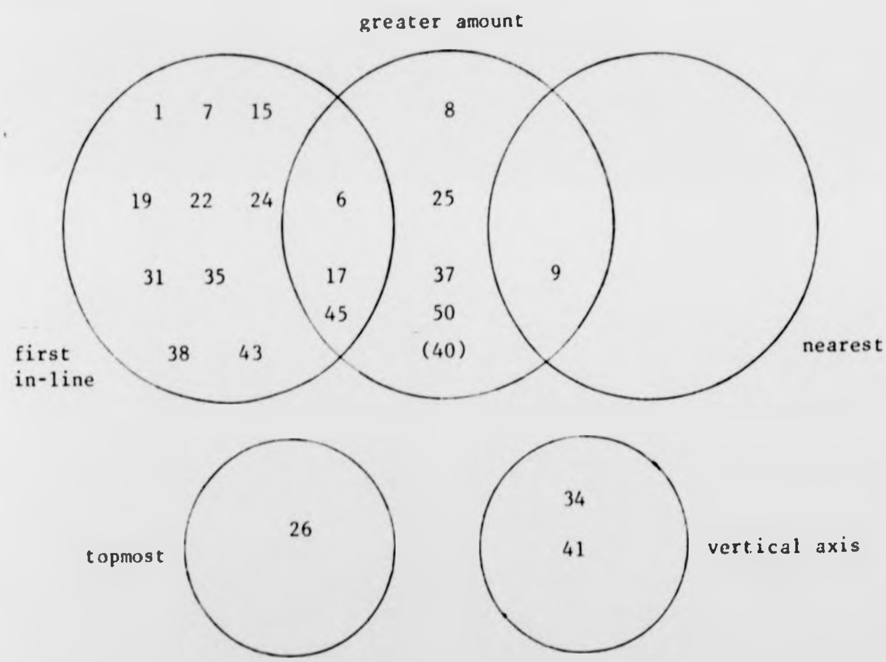


Fig. 16a Errors on Form A: H_8 - H_{12}

Form B

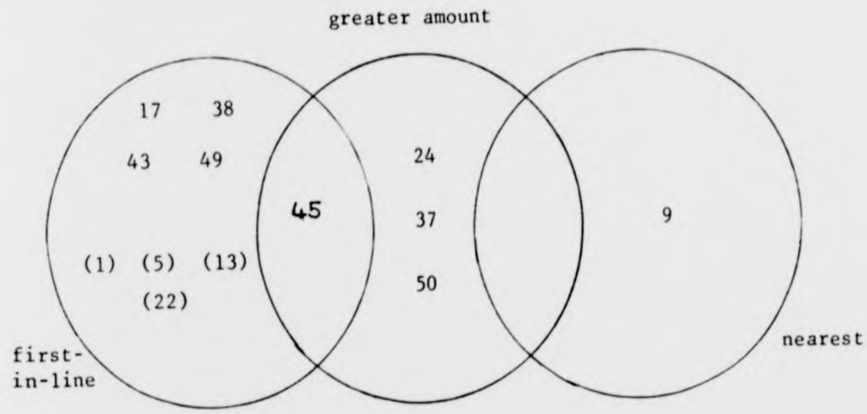


Fig. 16b Errors on Form B: H_8 - H_{12}

13.00 Answers to the questions asked in the small investigation of 30 children's preferences and interpretations of pictures in the BTBC Form A items 4, 20, 24, 29, 30, 31, 32, 47 and Form B item 31 are tabulated below. Where appropriate, the tables also show the results of chi-square one-sample tests. The results are presented in the order in which the questions were asked. (See Chapter 8, iii, 7.02).

13.01 Form A item 4.

Preferred toy	Age: 3	4	5	6	7	N
doll	1	0	3	2	1	7
ball	3	5	1	2	2	13
bat	2	1	2	2	3	10
						30

$\chi^2 = 1.8$ df 2; NS

Table 61 Preferences expressed for toys in item 4 (first choices).

	Age:	3	4	5	6	7	N	%
doll	identified correctly	1	1	3	3	2	10	33.3
	identified as 'girl'	4	4	3	1	4	16	53.3
	identified as 'lady'	1	1	0	1	0	3	10
	identified as 'person'	0	0	0	1	0	1	3.3
							30	
ball	identified correctly	6	5	6	6	6	29	99.7
	identified as 'a drum'	0	1	0	0	0	1	3.3
							30	
bat	identified correctly	1	4	5	5	5	20	66.7
	identified as 'a nail'	0	2	0	0	0	2	6.7
	identified as 'a screw'	1	0	0	0	0	1	3.3
	don't know response	4	0	1	1	1	7	23.3
							30	

Table 62 Identification of toys in item 4.

13.02 Form A item 20.

		Ages: 3	4	5	6	7	N	%
boat	identified correctly	6	6	6	6	6	30	100
doll	identified correctly	4	5	6	6	5	26	86.7
	identified as 'a girl'	2	1	0	0	1	4	13.3
							30	
bat & ball identified correctly		3	2	5	6	6	22	73.3
ball only identified correctly		3	2	0	0	0	5	16.7
'don't know' response		0	2	1	0	0	3	10
							30	

Table 63 Identification of toys in item 20.

		Ages: 3	4	5	6	7	N	%
Front	located correctly	4	4	5	4	4	21	70
	located at side*	1	1	0	2	2	6	20
	located behind	1	1	0	0	0	2	6.7
	'don't know'						1	
							30	
Behind	located correctly	5	5	6	6	6	28	93.3
	located in front	1	0	0	0	0	1	3.3
	located at side	0	1	0	0	0	1	3.3
							30	

Table 64A Location of 'front' and 'behind' in item 20.

* Several children remarked: "The doll/girl's at the front (i.e. at the front of the picture)

Preferred toy	Ages: 3	4	5	6	7	N
boat	2	2	3	2	5	14
doll	1	3	1	2	0	7
bat &/or ball	2	1	2	2	1	8
						29
settee	1					1

$\chi^2 = 2.95$ df 2 ; NS

Table 64B Preference expressed for toys in item 20 (first choices).

13.03 Form A item 24.

	No. of children who said jar had contents						
	Age: 3	4	5	6	7	N correct	% correct
Jar 1	5	5	5	4	6	25	83.3
Jar 2	1	1	3	4	3*	12	40
Jar 3	6	5	5	5	6	27	90

Table 65 No. of children correctly identifying jars with contents (item 24).

* Each of these 7 year olds at first said 'no' and then spontaneously corrected themselves.

13.04 Form A item 30.

No. of children who prefer each food

	Age: 3	4	5	6	7	N
ice-cream	4	4	3	2	3	16
pear	0	0	2	1	1	4
pie	2	2	1	3	2	10
						30

$\chi^2 = 7.2$ df 2; $p < .05$

Table 66 Preferences expressed for foods in item 30 (first choices).

No. of children who consider each food 'a pudding'

	Age: 3	4	5	6	7	N	%
ice-cream	4	4	4	2	2	16	53.3
pear	4	4	4	2	0	14	46.7
pie	6	6	6	6	6	30	100

Table 67 Identification of foods in item 30 as puddings.

13.05 Form A item 31.

Preferred shape	Age: 3	4	5	6	7	N
square	0	1	1	2	0	4
triangle	3	0	1	0	0	4
pentagon	2	1	1	1	2	7
circles	1	3	3	3	4	14
						29
'don't know'	0	1	0	0	0	1

$\chi^2 = 9.00$ df 3; $p < .05$

Table 68 Preferences expressed for shapes in item 31 (first choices).

Shapes judged 'alike'	Age: 3	4	5	6	7	N	%
2 circles	3	3	4	4	5	19	63
circle and triangle	1	0	0	0	0	1	3.3
circle and pentagon	0	0	1	1	0	2	6.7
circle and square	0	1	0	0	1	2	6.7
triangle and pentagon	0	0	0	1	0	1	3.3
square and pentagon	0	0	1	0	0	1	3.3
single shape only	2	0	0	0	0	2	6.7
don't know/'like what?'	0	2	0	0	0	2	6.7

Table 69 Responses to item 31.

Reasons given by the children for their incorrect responses

Circle and square

Pasquale, aged 7 : because that fits in that because it hasn't got sharp edges.

Vicky, aged 4 : This one's nice and this one's nicer.

Triangle and pentagon

Hena, aged 6 : because they've got those corners.

Square and pentagon

Emily, aged 5 : because it's like a square.

Two three year olds, one five year old and a six year old were unable to give reasons for their choices.

	No. of children who named the shapes correctly						
Age:	3	4	5	6	7	N	%
square	5	5	6	6	6	28	93.3
triangle	3	3	5	6	6	23	76.7
pentagon	0	0	0	0	0	0	
circle/round/ball/hoop	4	5	6	6	6	27	90

Table 70 Identification of 4 geometric shapes (item 31).

13.06 Form A item 32.

	No. of children favouring each car					
Age:	3	4	5	6	7	N
Car 1	2	1	1	0	0	4
Car 2	0	1	3	2	4	10
Car 3	4	4	2	4	2	16
						30

$\chi^2 = 7.2$ df 2; $p < .05$

Table 71 Preferences expressed for cars in item 32.

Reasons given by children for their chosen car

Car 3 Variants of "It hasn't got a roof" were given by 10 children; one of these also cited the spoked wheels. Two children produced "It goes faster". Two 3 year olds justified their choices by "Because I do" and "Because it is". Another 3 year old said "The sports car" whilst Rachel, aged 3, side-stepped the question "Why do you like that car best?" by saying "I don't know what colours they are. I like coloured pictures best."

Car 2 Four children gave variants of "It's going fast". One cited the boot ("It's got a thing that goes down at the back"). One five year old claimed "It's the biggest". Two children gave as the reason "It's in the middle" whilst John, aged 7, said: "That's difficult because they're all nice. That one looks like a racing car a bit" (car 2). A 4 year old could give no reason.

Car 1 The four children who selected Car 1 replied as follows:

Emily, aged 3 : It's going in there. (the tunnel)

Justin, aged 3 : Because it is.

Peter, aged 4 : It's got some funny wheels because they're under-over a little bit.

Ruth, aged 5 : Because that's the fastest car.

13.07 Form A item 47.

No. of children who chose each picture

Age:	3	4	5	6	7	N
Picture 1	2	2	4	3	2	13
Picture 2	3	2	1	2	0	8
Picture 3	1	2	1	1	4	9
						30

$\chi^2 = 1.4$ df 2; NS

Table 72 Preferences expressed for each picture in item 47 (first choices).

Reasons given by children for choosing each lollipop picture

Picture 1 Six children claimed that the first picture had the most lollipops including Ruth, aged 5, who said "There's five and you can give one to your brother and sister and Mummy and Daddy and one for myself". Two five year olds compared picture 1 with picture 3 but, whilst one noticed the different arrangement ("It's a different way from that") the other chose picture 1 "because it's the same as that, one, two, three, four, five." A six year old and a three year old gave no reason whilst Rachel, aged 3, avoided the question once more by saying "We've got some lollies at home" and Vicky's (aged 4) reason was "Because they're nicer".

<u>Picture 2</u> Three children chose the second picture "because they've just got three lollies"; Hena, aged 6, explained: "I like three lollies because I'd eat three but if there were lots I wouldn't". Two children gave no reason whilst Vincent's, aged 4, was "because they look nice." Christian, aged 5, said "They've got rounds inside" and Richard, aged 3, said "Them two, 'cos they're snail ones".

<u>Picture 3</u> One 7 year old didn't know why he chose the third picture and a 3 year old and a 4 year old each said " 'cos I do". Two children chose it because it contained five lollipops whilst Emma, aged 5, alleged that "them sticks are tallest" and Pasquale, aged 7, stated "They're bigger". Claire, aged 7, said "Because the've got those circles on and they're round and smaller than those." John, aged 7, chose by flavour: "They're blackberry. Those are strawberry and those are orange".

13.08 Form A item 29.

Age:	3	4	5	6	7	N
Squirrel 1 climbing or about to climb	6	6	6	6	6	30
Squirrel 2 climbing	2	4	4	1	3	14
doing something else	4	2	2	5	3	16
Squirrel 3 climbing	6	6	6	6	6	30

Table 73 Interpretation of squirrels' actions in item 29.

13.09 Form B item 31.

		No. of children					
Socks	Age:	3	4	5	6	7	N
striped		3	3	1	0	2	9
white		0	0	2	2	2	6
with clocks		1	0	2	2	2	7
patterned		2	3	1	2	0	8
							30

$\chi^2 = 0.66$ df 3; NS

Table 74 Preferences expressed for socks in item 31 (first choices).

CHAPTER 10

Discussion of the Results of the Standardisation of the BTBC.

The Sample

The standardisation sample appears to be representative of children in England in terms of urban/rural and social class distributions. The distribution of schools across rural, inner city and 'other urban' areas parallels the distribution found by HMI in its survey of primary schools (DES, 1978). The distribution of social classes shows a trend similar to the national distribution. (See Chapter 8, i, 2.07 and 2.15 - 2.18)

A sample restricted to northern England is not regarded as a serious limitation. Differences in reading performance between North and South, found in the APU surveys, were not significant though the performance of pupils in the metropolitan areas in the Midlands was significantly below that of pupils in metropolitan areas in the North and South of England (Gorman, White, Orchard and Tate, 1982). The mean score of the total sample of Northern pupils lies midway between that of the pupils from the South and the Midlands (54.6, 55.1 and 54.1 respectively). Associative relationships were observed between performance and background variables such as free school meals (ibid.). The patterns of performance between the three English regions appears to vary in relation to particular sub-categories of mathematics. However, "the variations in pupils' scores within regions is greater than between regions" (Foxman, Ruddock, Badger and Martini, 1982).

The schools used for the standardisation are limited to the maintained sector but, as stated earlier (Chapter 8, i) very few children under eight in the region attend private schools.

The absence of data on under-fives who do not attend nursery or pre-fives class might imply that the standard scores for three and four year olds should be treated cautiously. However, as homogeneity of age distribution was confirmed across SES groups, it can be assumed that the under-fives in the sample represent the distribution of social classes in the total sample population.

It would appear then that this English standardisation is considerably more representative of the nation's children than is Boehm's sample which is acknowledged not to be nationally representative of the U.S.A. (Boehm, 1971) and is certainly not representative of Great Britain.

Testing Procedures

It is claimed that the testing procedures adopted (Chapter 8, iii, 1.06 - 1.10) were satisfactory as they allowed each child to respond at his or her own pace, eliminated copying and were non-threatening. As all testing was done by the investigator there was no fudging of procedures or results by other testers nor were the results confounded by children who might have had difficulty in placing crosses as required by group testing.

Standard Scores

If the mean raw scores (Form A) of 5, 6 and 7 year olds are compared with those of kindergarten, first and second grade pupils tested at the beginning of the school year in Boehm's (1971) norming sample, it can be seen (Table 75) that the English five year olds from all SES groups and the six and seven year olds in SES group 3 obtain higher mean scores than their American counterparts (though the score differences between the two samples are small for middle and high SES kindergarten and 5 year old groups); whereas the English six and seven year olds in SES groups 1 and 2 obtain lower scores. It is not possible to say whether these are real differences in performance or whether they are because like is not being compared with like or because of the admittedly non-representative American sample.

	Socioeconomic Level		
	SES 3 / low	SES 2 / middle	SES 1 / high
age 5	28.5	32.5	36.3
kindergarten	25.5	31.8	35.8
age 6	36.8	38.8	41.6
Grade 1	33.8	42.0	46.3
age 7	41.6	44.2	46.3
Grade 2	41.2	46.9	47.3

Table 75 Mean raw scores of both samples by age or grade and SES.*

* See Tables 16 and 44; also Boehm, 1971 p.26.

The ceiling effect apparent in Figs. 3h, 3i, 4h and 4i is also observable in Boehm's sample. It suggests that with seven year olds the BTBC is only useful for identifying children who are behind their peers in their knowledge of the 50 BTBC concepts or individual concepts with which some children may not be familiar.

The accelerations, between 3:6 and 4:0 and between 4:6 and 5:0, observable in the standardisation curve (Fig. 5), are likely to be the result of new experiences in nursery and reception class. This may reflect the introduction of new vocabulary or it may be that the child is learning the skills of test respondent; that is, nursery or school has increased his ability and willingness to pay close attention to the words of the teacher or tester and improved his skill in the interpretation of pictures.

Sex Differences

The prediction and finding that there are no overall sex differences in the BTBC scores is in accord with current opinion and a number of studies (Macauley, 1977; Nelson and Bonvillian, 1978). There appear to be small differences of two to three points at 5½ and 7 years but, because for all other ages the differences are very small and these two are in opposite directions, this is assumed to be the result of sampling fluctuation and of no educational or psychological importance.

"The truth is that measures of linguistic proficiency, particularly for young children, are extrememly crude instruments... and thus it is not surprising that samples of linguistic behavior will reveal occasional differences between sub-groups... Probably several statistically significant differences in linguistic development could have been established if the samples of children had been divided into two groups on some arbitrary basis such as the initial letter of last name." (Macauley, 1977 p.357)

First and Second Half-year Testing

As no significant differences were found between the scaled BTBC scores of children tested during the first half of their school year and those tested during their second half separate norms have not been computed for the English standardisation. The BTBC (1971) manual gives separate norms for these two groups, probably to satisfy the requirements of U.S. legislation and practice in testing.

Inspection of Table 3 in the manual (App. 4, p.20) reveals that within Boehm's standardisation sample the children used to establish 'Beginning of Year' norms were not drawn from the same population as the children who were used for the 'Midyear' norming. Not only were the two groups drawn from different cities but the socio-economic composition of each group was different. Table 76 below provides a comparison of the two groups.

Socio-economic level		Beginning of year N	Beginning of year Percentage	Midyear N	Midyear Percentage
Kindergarten					
	Low	1921	54.6	162	18.7
	Middle	912	25.9	453	52.4
	High	684	19.4	250	28.9
	Total	3517		865	
Grade 1					
	Low	2303	49.4	276	27.9
	Middle	1313	28.2	435	43.9
	High	1043	22.4	280	28.3
	Total	4659		991	
Grade 2					
	Low	824	52.8	222	27.3
	Middle	381	24.4	349	42.9
	High	356	22.8	242	29.8
	Total	1561		813	

Table 76 Comparison of SES composition of Boehm's Beginning of Year with Midyear samples.

Moreover, the socio-economic groupings in Boehm's sample are based on the primary area from which each school population is drawn, not on the classification of each pupil.

However, notwithstanding the discrepancies between Boehm's two sub-samples and the vagueness of her SES classification, the midyear testing yielded higher mean scores within each grade and SES group (Boehm, 1971 p.26). Why were similar differences not found in the English sample? Two reasons are hazarded: firstly, Boehm's 'Beginning of Year' testing was done in September and October and the 'Midyear' testing between mid-November and late February whilst the English sample's first half-year testing was between September and mid- to late February; this period subsumes both the American periods. Possibly acceleration of concept learning or naming (or of skills in test responding) occurs after the first few months of the school year and then develops more gradually. If so, then this acceleration would fall within the first half-year of testing the English sample.

Secondly, the flexible admissions policies of British schools and the variety of organisation of age groups and classes within the Primary sector together with the preponderance of methods and pace geared to the individual child may not be commensurate with predominating practices in the American grade system.

Reliability Estimates

The mean KR_{20} coefficient of .81 and the mean standard error of measurement of 2.7 give confidence in consistency of scoring on the present adaptation and the standardisation of the BTBC (Form A). As explained earlier, the KR_{20} coefficients are likely to be slightly lower than the split-half estimates presented by Boehm (1971). The range of the reliability coefficients (.74 to .85) therfore compares favourably with Boehm's split-half coefficients (.68 to .90). As within Boehm's sample, the lowest reliability values are found with the oldest children, no doubt because of ceiling effects leading to a reduction in the variance of scores.

Short Tests C and D

The 25-item test C is seen as more satisfactory than the 20-item test D as a measure yielding results similar to those of the 50 item BTBC. The obtained correlation coefficients between scores on the short tests and the whole test (.94 Test C and .89 Test D) support the recommendation of using Test C, instead of the whole test, with preschool children and with some five year olds. However, because all the items in Test D are contained in Booklet 1 of the current edition, this may be, at present the most convenient short test to use. The advantages of both short tests lie in their brevity: maintaining the concentration of many preschool children throughout 50 items involves considerable skill on the part of the adult; after successive items which are far too difficult children often resort to guessing or show signs of inattention. Moreover, as will be made clear in the next chapter, certain items produce correct responses from the nursery children when the concept is not understood.

Alternate Form Equivalence

The mean correlation coefficient .80 observed between the Form A and Form B scores of 144 children and the median coefficient of .815 (Table 32) compare favourably with the AB coefficients observed by Boehm (median .76). Thus the results of the present comparison of Forms A and B confirm Boehm's conclusion that "the two forms yield essentially equivalent raw scores" (Boehm, 1971 p.21).

However, even though the total scores on each form are comparable, certain items are not of equal difficulty on both forms. The content of these items will be examined in the next chapter.

Social Class Differences

The highly significant differences between the mean scaled scores of the three broad SES groups 1, 2 and 3 bear out predictions. Numerous British studies have found strong links between social class and scholastic achievement (e.g. Douglas, 1964; Wiseman, 1967; Goodacre, 1968; Brandis and Henderson, 1970; Davie, Butler and Goldstein, 1972).

The results do not support the prediction that schooling might lessen the gap between the mean scores of SES group 3 and children higher in the social scale. The mean scores of SES group 3 remain low throughout the infant school years whilst those of the middle SES group hover around the sample mean and the mean scores of the high SES group remain well above average.

It must be emphasised that these results are indicative only of group trends and that the considerable overlap of scores across the groups means that socio-economic status cannot efficiently predict the score of any child. The complex web of variables associated with social factors, such as the quality of housing, the educational level of parents and child-rearing practices, produces enormous variety of achievement within any one social class. For example, one six and a half year old who achieved a BTBC raw score of 50 (T score 72) is the eldest child of a trucker and a young mother who live in a "two-up, two-down" inner-city slum dwelling without indoor sanitation.

A random sample of 45 children (5 at each age level) illustrates the social class trends noted above as well as the variations (See App. 15). Table 77 below shows the distribution of these 45 children's scaled scores by SES grouping and school location.

SES group	Rural	Urban	Inner City	Rural Disadvan.	Urban Disadvan.	Total N
SES 1	64 59 59 56 46	52 48	59			8
SES 2	66 54 48 46	63 56 55 54 44 44 41 42		51 46	54 36	16
SES 3	63 62 52	54 37 39	68 48	44 43 26	48 37	13
SES 4 (group 2 or 3)	48					1
SES 5 (group 1 or 2)						0
SES 6 (group ?)	53 52 45 45 41			40	41	7
N	18	13	3	6	5	45

Table 77 T scores of 45 children by SES and location.

Of the four children (in this sub-sample) in SES group 3 who obtained scores greater than 52 two have unemployed single parents, one has both parents unemployed whilst the fourth attends a school (40) which surprised the tester by the high level of performance of most of the children. (The highest scorer of this sub-sample also attends this school.) It would be foolish to generalise from this small sub-sample but these cases illustrate the pitfalls that can arise through crude SES classifications. Single parent families and the unemployed are the most obvious examples; split or bereaved or jobless families may be classified correctly according to the Registrar-General but this may be a poor indicator of family mores and aspirations. For this reason, only families who had been unemployed for a number of years were classified as unemployed for this research. The massive scale of unemployment in recent years is likely to have upset the niceties of the Registrar-General's scale in so far as this is used to relate level of occupation to other variables. The breadwinner made redundant often has only the option of a job below the level of the previous one or the dole whilst young people may fail to enter the occupation for which they are qualified.

Therefore, to produce three sets of norms for high, middle and low SES groups, as presented in the BTBC manual, would be inappropriate. Even if it were feasible it is considered undesirable. Boehm's instructions for obtaining a child's percentile grade require the teacher or psychologist to "locate the child's raw score in the appropriate column for grade and socioeconomic level... ...a child in kindergarten at a middle-socioeconomic-level school..." (Boehm, 1971 p.26) (underlining added). (Mixed catchment areas are disregarded by Boehm.) The dangers of labelling schools, groups or individuals, so that expectations of achievement of those at the bottom of the ladder are lowered, have been well argued (Pilling and Pringle, 1978; Pringle, 1974; Nash, 1973; Rosen, 1972; Bernstein, 1971, 1970; Pidgeon, 1970; Rosenthal and Jacobson, 1968).

In any event, it appears that a child's social class is less indicative of school success than are parental attitudes and aspirations and styles of parenting (Light, 1983; Brophy, 1970; Wiseman, 1967).

Location of School and Home

The results suggest that rural children tend to show higher performance on the BTBC than either inner city or other urban children (Table 45). The APU surveys observed higher mean reading scores of ten year olds in non-metropolitan than in metropolitan areas (Gorman et al., 1982). It is uncertain what factors contributed to the superior performance of either of these rural samples. There are likely to be a multiplicity of both home and school variables.

When the BTBC scores of the same SES groups are compared (Tables 47 and 48) the results suggest that children of the same social class tend to get higher scores on the BTBC if they go to school in a rural locality, that is not notably marked by social disadvantage, than if they attend a school in any other type of area. Children who live and go to school in urban rehousing estates tend to perform less well than their counterparts in the inner city or in other urban areas or in non-disadvantaged rural areas.

The disadvantaged urban areas in the research sample, outside the city or town centres, happened to be located in six estates into which families had originally moved from a city or town slum area. The oldest estate dates from the late thirties with post-war extensions; building began on the others in the fifties and sixties. So each of these estates contains a mix of families from first and subsequent waves of rehousing. At least one estate is obviously used as the dumping ground for the housing department's problem tenants. The smallest has become an isolated drab community which, according to community workers and the school, lacks cohesion and is marked by constant bickering between neighbours which erupts into serious violence.

It is not implied that the other urban areas do not contain pockets of social disadvantage but their school populations consist of a mix of social backgrounds, some preponderantly middle-class, others preponderantly working-class, with varying proportions of private and council houses. Because of its high 'immigrant' population one of the latter group has amassed sufficient points to be graded SPA.

Similarly, the rural schools labelled 'non-disadvantaged' draw on very mixed catchment areas some of which contain small pockets of social disadvantage or isolated families; even the 'stockbroker belt' commuters' villages have families living in poverty. The squire's son, who will eventually go to prep and public school, sits alongside the labourer's daughter and the miner's son in the infant class of the village school.

Barry Hines portrayed Kes growing up in a community which typifies those surrounding the six socially disadvantaged rural schools. They are relatively isolated mining communities where the pit no longer provides full employment, where poverty is rife, where the quality of housing has deteriorated and where the social services are kept fully employed.

These are the sorts of communities which this research compares with the present-day inner city areas. The latter also have their share of social problems and there is no evidence that they have decreased since the time of the Plowden report or the EPA research studies of the early seventies (Halsey, 1972; Field, 1979; Jackson, 1980; Rutter, 1981; Gibson and Langstaff, 1982; Murie, 1983; Ward, 1977). Indeed the plight of the poorest has been exacerbated by the reduction in the supply of low-rent housing and by the acceleration of decay in doomed slum areas. Yet the mean BTBC score of these inner city children is significantly higher ($p < .02$) than that of children in the same SES groups who live on urban rehousing estates away from the centre of a conurbation. (Their mean score was also slightly higher than that of the rural disadvantaged children and barely one point lower than that of other urban children of the same socio-economic status but neither of these differences reached statistical significance.)

Why did these inner city children perform contrary to expectations? The social problems of a community's poorest inhabitants, as well as their feelings of powerlessness and failure and the consequent adoption of a survival strategy of adaptive retrenchment (Wilson and Herbert, 1978), are common to most urban environments where poverty (or relative poverty) is rife. Children who have ever experienced poor housing conditions still have lower attainments by the end of their compulsory school years regardless of their current socio-economic conditions (Essen and Wedge, 1982). The families on the estates removed from the city centre also experience the added difficulties associated with isolation from the extended family and city life, the segregation of the generations, increased fuel and transport costs and the lack of employment for women; all of which affect the mother of small children tied to her immediate environs. So perforce a view of the world dominated by mistrust and lack of curiosity (Wilson and Herbert, op. cit.) is entrenched more deeply and reflected in the socialization of the children.

Variables such as these can only be viewed as possible contributory factors to the relative success of the children who remain in the inner city. Two other sets of variables are likely to be of major importance. One set relates to the social composition of the present-day inner city and the other to the schools.

Newly divorced or separated mothers often move into the run-down parts of the city: some to be near relatives, others because it is the only place where they can find cheap, rented accommodation. The present socio-economic status of such parents may not, therefore, indicate the level of their education or aspirations. There may now be more of a social mix in some inner city schools than was the case a decade or more ago.

It is likely that the relative success of these inner city children is, to a large extent, attributable to their schools. With hindsight it is regretted that systematic observation of school variables was not designed at the outset of this research. The observations which follow are therefore largely speculative but it is posited that they are germane to the present-day picture of these particular inner city schools in contrast to the pictures painted by earlier studies (e.g. Halsey, 1972) and by the Plowden report.

Firstly, four of the six 'school adversity' features found to be strongly associated with children's problems (Rutter, 1977) did not apply to this sample: Hull does not have a high 'immigrant' population; absenteeism was not noticeably more common than in schools in other types of areas; the pupil/teacher ratio (excluding nurseries) was approximately 19.3[1] and the average class size was approximately 22.8. The high teacher-turnover, common in the past, is no longer a feature of these particular schools. The highest incidence of staff changes and absenteeism was observed in a run-down estate on the outskirts of a town.

Secondly, it is surmised that, after the Plowden report drew attention to the needs of inner city schools, not only were material resources increased but the plight of these under-privileged children captured the sympathies of many young teachers who rose to the challenges which they presented. Only the more able teachers were likely to survive the harsh reality of life in an EPA school whilst their weaker brethren escaped to the safety of the suburbs. Education authorities were concerned to tackle the problems of inner city schools and, as the teacher shortage declined, were more selective in teaching appointments to these schools.

[1] In January, 1980 the pupil/teacher ratio (including headteachers) in Primary schools was 22.9 in Barnsley, 21.6 in Doncaster and 22.4 in Humberside. The average ratio was 22.6 for all English metropolitan districts and 23.3 for non-metropolitan counties. (DES, 1980)

Thirdly, almost all these children now have access to at least part-time nursery education and those with identifiable problems can attend full-time.

Fourthly, during the seventies, sociolinguistic research focussed attention on the language of young disadvantaged children. Projects such as "Communication in Early Childhood" (Tough, 197) involved nursery and infant teachers. It may be that their concentration on children's language and their efforts to improve performance have achieved some success in these inner city schools. Perhaps the key factor is the quality of the teaching in these schools in recent years.

Observation of classroom teaching was not part of the present research. Nevertheless, it was impossible not to catch the ambience of the schools visited: the staff of these inner city schools, though often working in appalling conditions, gave evidence of (1) thorough knowledge of their children, (2) a welcoming and concerned approach to parents, (3) interest in the neighbourhood and its problems, (4) cheerfulness, resilience and humour, (5) excellent relationships with the children. (Only one teacher was encountered who belied this.) Apart from one rare exception it was obvious in each school that the headteacher and staff worked well as a team.

The worst building of all exemplifies the description of EPA schools given in paragraph 133 of the Plowden report. However, the grim exterior and its environs are counteracted by a large notice inside the entrance saying "Welcome to — ", by the gusto of the dynamic headteacher and supporting staff, by the bustle of active and friendly children, by the obvious friendly relationships with parents and by the liveliness in the tiny, cramped staffroom.

It is not known how far conditions in the nation's inner city Primary schools, or the achievements of their pupils, have altered since the Plowden report. It may be that the present results and observations are a-typical. Cities are not all alike in having high rates of problems (Rutter, 1981) and there are wide variations in the quality of inner city education (Little, 1977). Notwithstanding the limitations of the foregoing observations, this sample of inner city schools lends support to arguments for the regeneration of inner city life rather than wholesale removal of its inhabitants to the urban deserts of new estates. It also highlights the need to update information on life in our inner city Primary schools.

CHAPTER 11

The Validity of the BTBC

"A test's validity is the extent to which scientifically valuable or practically useful inferences can be drawn from the scores" (Jensen, 1980, p.297). This chapter attempts an assessment of the inferences that may be drawn from the BTBC scores.

After an analysis of each BTBC item the content and predictive validity evidence provided by the results of the investigation are summarised. Then follows an examination of the test's constructs. Finally, the present findings in relation to Eve Clark's theory and non-linguistic strategies are considered.

Analysis of the BTBC items

Each item in turn is discussed in detail below. The main considerations are firstly, whether a child who passes an item does so because he understands the concept and its label or whether he could pass (guesswork apart) without any, or only partial, understanding. Secondly, predictions are made about whether a child who understands the concept might fail the item.

Reference is made to the content validity study (Chapter 8, 4.00 - 6.00 and Chapter 9, 8.00 - 8.16), the form equivalence study (Chapter 8, 2.00 - 2.02 and Chapter 9, 6.00) and the small investigation of children's interpretations of the drawings (Chapter8, 7.00 - 7.02 and Chapter 9, 13.00 - 13.08). The latter is referred to as the 'picture' sample. When comparing Forms A and B the overall facility values of each Form may be quoted; these are based on the sub-sample (N = 144) in the form equivalence study. Otherwise the facility and discrimination values cited are those from the total sample (N = 1928) refer to Form A only and are for each age level. They can be found in Table 18.

Errors on Form A were inspected for the total sample but only the errors of 144 subjects are available for Form B. Thus differences between errors on Form B may be numerically small and non-significant even when the same item on Form A produces highly significant differences in errors. (See Chapter 9, section 12; Appendix 16 provides an index of which tables contain information on errors on each item.)

Various possible reasons are advanced for the patterns of success and failure on each item. The possibility of arriving at the right answer by guessing applies to all the 50 items: this will not be repeated fifty times. So in each case the possibility of achieving success through guesswork must be taken as read.

Item 1: top

Though this item is very easy for infant school children, and therefore discriminates poorly from 4 years upwards, it is satisfactory as a first item on both Form A and Form B. Content validity results show that it is a concept with which all the age groups (3:6 - 7:6) are familiar and therefore gives even the youngest children the notion that the 'game' or 'puzzle' is something they can do.

The facility values appear to support Clark's (1973, 1977) hypothesis that terms which orient to the upper end of a vertical axis are acquired early. It is, however, possible that, because of the salience of top, some of the youngest children might pass this item without a complete understanding of at the top. The top or upper sides predominate in three year olds' choices regardless of the orientational term used in instructions (Clark, 1980). Therefore, as Clark points out, semantic knowledge should not be attributed to the child until he contrasts at least one other term with top. Conversely, a child who fails item 1 can be assumed to have less than complete knowledge of top.

Errors suggest that the majority of children with an incomplete knowledge of top do not confuse it with bottom or, at any rate for most of them, a first-in-line strategy may override this confusion: on Form A over two thirds of wrong choices were of the first picture ($p < .001$).

Item 2: through

Item 2 is also satisfactory as an introductory item. It cannot be assumed, however, that all children who pass necessarily understand through: some may have applied a first-in-line strategy. Not all children who fail necessarily lack comprehension of through: on Form A the passive voice may confuse some children. The preponderance of the second picture in wrong responses ($p < .001$) suggests that "the bead that has a string through it" may have been interpreted as "the bead through the string". It is also possible that ambiguity in the meaning of going (Form B) may lead some children to choose the wrong picture i.e. either of the two which depict the dog going to jump through the hoop.

Through features in all the curriculum areas examined.

Item 3: away from

Content validity results do not suggest that this phrase is crucial to understanding teachers' directions in the early years at school. It was not observed in any of the nursery or reception classes. Occasionally an older child would be told to move away from x. It seldom features in Mathematics workbooks. Nursery and infant school children associate away with 'put/clear away'; which may account for this item being found easier on Form B than on Form A where the box under the table may be considered to be 'put away'. However, this probably means that those children who get the item right have understood the term. Boehm (1971) also finds more children succeeding on Form B.

Item 3 discriminates best below 5½ years. A comparison of the facility values of items 3 and 4 support the semantic features hypothesis (Clark, 1973): young children find the positive term next to considerably easier than the negative away from. Away from is generally well-understood by infant school children but elicitation tasks have been found to produce a low use of away from and other words which describe the opposite of nearness such as far (Durkin, 1978). There is an overwhelming tendency in children to perceive the location of an object in terms of its nearness to other objects or oneself (ibid.).

Item 4: next to

This term was observed in teachers' talk with all the age groups and in Mathematics and Science schemes though it seldom appears in reading books but next (spatially and temporally) was common.

When four to seven year olds are asked to place one object next to another the youngest children put the objects in contact; the older the child the greater is the distance from the reference point (Durkin, 1978). However, this concept is context dependent: even the youngest children left spaces between toy cars because cars do not normally touch (ibid.). So if the youngest children respond correctly, to item 4 on Form A where the reference point is a truck, it cannot be taken for granted that they accept that something next to another object does not have to touch it. (Item 4, Form B consists entirely of animals.) Data from the Form equivalence study (N = 144) shows five children who failed item 4 on Form A and passed Form B and five who failed Form B and succeeded on Form A. Piaget (1956) demonstrates that young children's spatial ideas rest on primitive topological relationships: things are together or separated. A person on a see-saw is next to the see-saw. The perceptual salience of nearness may allow some children to pass item 4 with an incomplete understanding of next to.

The drawings on Form A may lead to confusion: only a third of the 'picture' sample identified the doll correctly; two thirds identified it variously as a girl, lady or person. Moreover, a third of the children could not identify the bat. Thus for some children the only toy seen in the picture is the ball. Hence some children may experience conflict when directed to point to the toy next to the truck.

Item 4, Form B may also present ambiguity to the youngest children: it is based on the assumption that a rabbit, cat, monkey and dog are already classified as animals; whereas some younger children may not have learnt this classification (Anglin, 1977; Blewitt, 1983).

So, although item 4 is easy for children over 4 years, when the BTBC is used diagnostically the results from this item need further investigation.

Item 5: inside

The ease with which this item is tackled by all ages probably reflects the young child's non-linguistic preference for putting things in containers (Clark, 1980; Wilcox and Palermo, 1974). Inside appears in all the curricula materials examined though in is more frequently used.

A first-in-line strategy could possibly account for a few children passing item 5 on Form A but as 92% of the sub-sample made the same score on each of the Forms it seems unlikely that this is more than a rare occurrence. As the tester clearly enunciated the first syllable, failure to locate inside can be assumed to signal an incomplete understanding of the concept or of the label.

Item 6: some but not many

Though the four concepts in this composite item appear in curricula materials and in teachers' discourse no instance was found of all the parts combined as in item 6 and only rare instances of not many. The Form A version is easier than Form B (F = .88 and .66 respectively) and was also found so by Boehm's (1971) sample. The explanation appears to be that it requires more sophisticated judgment to select the vase with some but not many flowers from three vases containing 3, 6 and 1 flowers than it does to select the box with some but not many marbles from three boxes containing 11, 0 and 4 marbles.

The wrong responses on each Form suggest at least three strategies which might have been employed. Over three times as many children chose 11 marbles rather than the empty box ($p < .001$). This may reflect young children's preference for the larger amount (Clark, 1977) or it may mean that some children understood the combined concept in terms of its opposite (which they could have interpreted as 'a lot') or they may simply have pointed to the first picture in the row. If children had used this last strategy it would have enabled them to answer Form B correctly without understanding the concept. The 'greater amount' strategy is barely apparent in the wrong responses to Form B: nine times as many children chose the single flower rather than six flowers.

It seems therefore that some children could pass item 6 on either Form without cognizance of the total combination of the concepts involved; others who grasp not many might fail Form B because of its poor design.

Words that contain 'not x' tend to be acquired later than their positive counterparts (Clark, 1977) and as young children frequently interpret only the main clause in a sentence (Townsend and Erb, 1975) it is likely that many of the younger children attended only to some. Alternatively, some children may have attended only to not many (or even just many) as these words are attached directly to the concrete noun. Cognizance of not many could produce the correct answer on Form A and the single flower response to Form B. The conjunction but seldom appears in the speech of three and four year olds (Smith, 1977) and therefore it may have been ignored by some children.

Item 7: middle

Middle appears in all the sampled areas of the curriculum though not as frequently as expected. The proportion of children who pass this item rises steeply during the nursery and reception class years and thereafter becomes a well-understood concept. Discrimination values decline with increasing age.

It is not known why 16 children failed Form B yet only four failed Form A (sub-sample N = 144). Because of this discrepancy the only conclusions that can be drawn from children's scores on item 7 are

(a) children who pass Form A can identify the middle position in this context;

(b) children who fail Form A do not have a complete knowledge of middle;

(c) children who pass Form B probably understand the concept of the middle position in an array;

(d) children who fail Form B, though unable to identify the middle in this context, may or may not be able to do so in other contexts.

On the other hand, Boehm's (1971) sample appears to have found Form B slightly easier than Form A.

Item 8: few

The Rasch analysis showed low scorers on the BTBC doing better than expected on item 8 whereas children whose raw score was above 36 did less well than expected. Facility values increase with age until 5 years; then a plateau occurs (F = .81) until $6\frac{1}{2}$ and 7 when the proportion passing the item drops to .76 and .79 rising to a peak of .85 at $7\frac{1}{2}$ years. Thus it appears that with increasing age a child's interpretation of the item content alters in some way. It could be that some sixes are in the process of acquiring the concept 'least' and for a time few and least become confused. The change in the pattern of wrong responses around 6 years seems to support this inference. Almost three quarters of the under-sixes who failed item 8 (Form A) chose the greatest number of cakes whereas only a third of the 6 and 7 year olds did so; the single cake is the wrong response for two thirds of these older children ($p < .001$ in each case).

Children who apply the first-in-line strategy could pass item 8 without understanding few but children who do understand are unlikely to fail.

Few appears in curriculum materials though not frequently.

Item 9: farthest

Farthest appears in "Science 5/13" but was not observed in any other curricula materials nor in teachers' discourse. However, far and farther feature in reading books and sometimes in teachers' speech. Facility and discrimination values suggest that this is a useful item.

The smaller proportion of success on Form B (also seen in Boehm's results) probably reflects the differences between item content on each Form: judging the boat farthest from the shore (Form A) depends on the use of size cues (Vernon, 1968) whereas there are no size cues to the dress which is farthest from the socks hanging on a washing-line (Form B). (On Form A the relative distances actually differ by about half a centimetre; on Form A they each differ by about 2cm.)

The drawing on Form B appears free from ambiguity. The drawing on Form A could mislead: without first pointing out the shore (a precaution not included in Boehm's test directions) there is no guarantee that children unfamiliar with a wooden jetty do not perceive the horizon line as the shore. The Humberside children tended to respond with greater alacrity to this item on Form A than did the Barnsley and Doncaster children.

It seems that on both Forms a child has to both understand the concept farthest and have knowledge of its label in order to pass item 9 though Form B provides the more stringent test of concept understanding in that farthest is not tied to the far distance.

Wrong responses to both Forms of item 9 show that when the concept or its label was not understood the nearest object was selected ($p < .001$). This could be an instance of the nearest object having greater perceptual saliency (Durkin, op. cit.) or it could be that for some children farthest may be partly understood in terms of its unmarked opposite or both tendencies may operate together. Additionally, in the case of Form A only, a preference for the larger object may also influence choice: the nearest boat also appears to be the largest one if one's comprehension of perspective is uncertain.

Item 10: around

For school age children the high pass rates give this item poor discrimination values; whilst for the youngest age group the very low discrimination value suggests that some children are getting this item right for the wrong reasons. It is suggested that this is another example of the greater amount (eight balls), which happens to be the right answer, predominating over lesser amounts (four balls). On Form B, where different amounts are not involved, the facility values are much lower (.67 as against .95 in the Form equivalence sub-sample). This difference was not found by Boehm (1971).

not included in Boehm's test directions) there is no guarantee that children unfamiliar with a wooden jetty do not perceive the horizon line as the shore. The Humberside children tended to respond with greater alacrity to this item on Form A than did the Barnsley and Doncaster children.

It seems that on both Forms a child has to both understand the concept farthest and have knowledge of its label in order to pass item 9 though Form B provides the more stringent test of concept understanding in that farthest is not tied to the far distance.

Wrong responses to both Forms of item 9 show that when the concept or its label was not understood the nearest object was selected ($p < .001$). This could be an instance of the nearest object having greater perceptual saliency (Durkin, op. cit.) or it could be that for some children farthest may be partly understood in terms of its unmarked opposite or both tendencies may operate together. Additionally, in the case of Form A only, a preference for the larger object may also influence choice: the nearest boat also appears to be the largest one if one's comprehension of perspective is uncertain.

Item 10: around

For school age children the high pass rates give this item poor discrimination values; whilst for the youngest age group the very low discrimination value suggests that some children are getting this item right for the wrong reasons. It is suggested that this is another example of the greater amount (eight balls), which happens to be the right answer, predominating over lesser amounts (four balls). On Form B, where different amounts are not involved, the facility values are much lower (.67 as against .95 in the Form equivalence sub-sample). This difference was not found by Boehm (1971).

Why 38 children, who failed item 10 Form B, chose the flower which is crossed by the string might be explained by reference to Clark's assertion (op. cit.) that certain terms are first understood in terms of their opposites. However, in this case, syntactical reasons are more plausible: some of these children may have confused the subject-verb-object pattern of the passive with the active voice and chosen the flower around the string. A similar pattern of responses happened with Form A but these differences were not statistically significant even though the balls in a box are a more natural arrangement than balls around a box. Thus item 10 appears to test the youngest children's syntactical knowledge as well as their understanding of around.

Though around appears in the infant school curriculum it was not observed in nurseries: round is generally used instead.

Item 11: over

From 4:6 upwards very few children have difficulty with this item. At 3:6 and 4:0 the reversal errors on Form B suggest that some children either confused over with under or perceived the crouching child as having fallen over!

When 2 and 3 year olds were asked to place an object over a fence Clark (1977) found that a third of those failing the task put the object under the fence and approximately half tried to balance it on top of the fence. It is by no means certain that children who pass item 11 can distinguish between over and on; under is not depicted on Form A and is ambiguous on Form B. Item 11 could be improved on each Form by redrawing with one object clearly under, one over and one on the surface of the referent.

Over appears in all the curriculum materials examined and in teachers' discourse though often with different connotations as in 'fall over', 'over

Why 38 children, who failed item 10 Form B, chose the flower which is crossed by the string might be explained by reference to Clark's assertion (op. cit.) that certain terms are first understood in terms of their opposites. However, in this case, syntactical reasons are more plausible: some of these children may have confused the subject-verb-object pattern of the passive with the active voice and chosen the flower around the string. A similar pattern of responses happened with Form A but these differences were not statistically significant even though the balls in a box are a more natural arrangement than balls around a box. Thus item 10 appears to test the youngest children's syntactical knowledge as well as their understanding of around.

Though around appears in the infant school curriculum it was not observed in nurseries: round is generally used instead.

Item 11: over

From 4:6 upwards very few children have difficulty with this item. At 3:6 and 4:0 the reversal errors on Form B suggest that some children either confused over with under or perceived the crouching child as having fallen over!

When 2 and 3 year olds were asked to place an object over a fence Clark (1977) found that a third of those failing the task put the object under the fence and approximately half tried to balance it on top of the fence. It is by no means certain that children who pass item 11 can distinguish between over and on; under is not depicted on Form A and is ambiguous on Form B. Item 11 could be improved on each Form by redrawing with one object clearly under, one over and one on the surface of the referent.

Over appears in all the curriculum materials examined and in teachers' discourse though often with different connotations as in 'fall over', 'over and over again', 'knock over', 'turn over' and 'change over'.

Item 12: widest

The word widest was never observed in teachers' speech and, with the exception of two of the Mathematics schemes' workbooks, is absent from all the curriculum materials examined. Wide and wider were observed more often.

Facility values range from .37 at 3:6 to .94 at 7:6. Discrimination values are highest at 4:0 and 4:6 (.44 and .30). The commonest error is for the narrowest door (Form A, $p < .001$) or tie (Form B, NS) to be chosen. Wide/narrow is one of the later pairs of dimensional adjectives to be acquired (Clark, 1977; Eilers et al., 1974). As the objects on each Form are not presented in order of size this item theoretically requires the child to compare and reorder; whether this is what actually happens, with the children who produce the right answer, is not known. It seems likely that children who fail do so because they have, at least, an incomplete understanding of widest but some children may pass by choosing the largest object. In addition, they could pass Form A by applying the first-in-line strategy. Slightly more children pass Form A than Form B and a quarter of the sub-sample scored differently on each Form. The youngest children in Boehm's (1971) sample were more successful on Form B than Form A.

Item 13: most

This appears to be a usefully discriminating item for preschool and reception class children; thereafter it becomes very easy for children of at least average ability. Most features in all the curriculum materials examined, excepting Bate and Smith (1978), though it appears far less frequently than more.

Item 13 was found to be slightly more difficult on Form B than on Form A. The box with the most eggs (Form A) has an egg in each compartment, thus conveying absolute fullness, whereas the arrangement of buttons (Form B) does not: there is still room to add more buttons. For children under five there is something compelling about "the notion of fullness so that it can override other criteria and powerfully influence the judgment of each truth value" (Donaldson and McGarrigle, 1974, p.186). In addition, supermarket egg boxes are more familiar to children than boxes of buttons.

So there is no guarantee that a child who passes item 13 on Form A is not confusing most with fullness. Neither is there any guarantee that any child who answers correctly on either Form is not merely responding by stating a preference for the greater amount. On the other hand, children who fail can be assumed to have, at least, an incomplete understanding of most.

Though the box with the least eggs or buttons was the most frequent error on each Form the differences between the wrong responses were slight and without significance. Nevertheless, it suggests that some children have acquired only partial meanings of most and least and that some of them recognise that most and least are on the same dimension. Studies have consistently demonstrated that more and less do not have their adult meanings for young children but there is disagreement about the ways in which the definitions differ. It has been suggested that the definition of more is overextended (Donaldson and Balfour, 1968; Donaldson and Wales, 1970; Clark, 1973, 1977). Alternatively, inappropriate responses to tasks involving less do not necessarily indicate that less is understood to be a synonym for more (Carey, 1976; Trehub and Abramovitch, 1978; Hudson et al., 1982). Although it appears that children acquire a mature understanding of more before less this is difficult to prove empirically: a test is required which differentiates between children with a full linguistic understanding of both more and less and those who have a full semantic knowledge of only one of these terms. When consistency of performance on less choices is compared with children's explanations of more choices which include the simultaneous comparative usage of more, there is some evidence that supports the notion of an overextension of more but the majority of wrong less choices appear to be based on non-linguistic information (Hudson et al., 1982 ; see also Trehub and Abramovitch, 1978 , Townsend and Erb, 1975 , Richards, 1979a, Clark, 1980 and Laxon, 1981).

It is thus apparent that the BTBC cannot provide unequivocal information about the maturity of a child's understanding of either most or least (item 50).

Item 14: between

Tryadic relations are perceptually more complex than dyadic ones. Not only is there more visual information to organise but there is also more linguistic complexity (Durkin, 1978; Sinha and Walkerdine, 1974). In the case of item 14 the language of the instruction is relatively simple. Most four to seven year olds have some semantic knowledge of between though responses to tasks, where a nonsense syllable is substituted for the label between, are poor (Durkin, op. cit.) Tasks involving between are exceptionally difficult for three year olds (Sinha and Walkerdine, op. cit.).

Responses to item 14 on both Forms show it to be context dependent: Form B pictures produce more errors despite the appearance of greater simplicity. It may be that, for some children, between is semantically easier to comprehend in a horizontal array (Form A) than in a vertical one (Form B).

Most of the errors on Form A show that the significant word in the direction is 'spoons': either the first spoon or the cup next to it was chosen most frequently ($p < .001$). Uniquely, one bright seven year old, whose total raw score was high, confidently pointed to both the cups and she did this again after the test when the item was repeated as a check.

As she was syntactically fluent it seems that she thought between meant its opposite.

In teachers' discourse between was only observed with seven year olds and its appearance in the curriculum materials is not frequent. Both facility and discrimination values point to its usefulness as a BTBC item with the best discrimination at 4½ and 5 years.

Success on item 14 probably indicates a mature comprehension of between but failure does not necessarily imply that there is no understanding of the concept. Durkin (1978) suggests that the linguistic complexities attendant on the use of between are the main source of difficulty that children face in the comprehension and use of this term. Quite young children comprehend the concept of between when the task is sufficiently simple (ibid.).

Item 15: whole

This item discriminates most usefully from 4:6 to 5:6. It is a difficult item for the youngest children and easy for most six and seven year olds.

Confusion of whole with hole may cause problems in interpretation. The most common wrong response to Form A ($p < .001$) was the first cake in the row. This may mean either that the non-linguistic strategy of choosing the first object predominated or that the cake with the largest hole was selected.

The correct answer to item 15 is also the greater amount and, additionally, on Form B it is the first object in the row. So there is the possibility that some children may respond correctly without understanding whole. Conversely, failure does not necessarily mean that the concept of 'wholeness' is absent.

Item 15 appears to be slightly easier on Form A than on Form B both in this sample and in Boehm's. This may be because the Form A drawing clearly depicts the whole cake whereas the Form B drawing may be seen by a child as a side-view of the apple.

The word whole was seldom observed in teachers' discourse; it appears in few Mathematics books and occasionally in only one of the reading schemes and does not feature at all in the two "Science 5/13" books. However, it is included in the manual for Assessment in Nursery Education.

Item 16: nearest

This item was passed easily by the vast majority of children under 5:6 and by the total sample aged from 5:6 to 7:6. Item 16 would make a useful first item on the test.

The word nearest appears only infrequently in curricula materials and was observed in teachers' speech only with the older children. Thus it seems that nearest is a concept learnt during the years before school. Along with next to (item 4) it is acquired earlier than its marked opposite farthest (item 9) and the term away from (item 3).

It seems that all the children who pass item 16 understand nearest and the few who fail still have difficulty with the concept or its label.

Item 17: second

As expected, this item discriminates best between 4:6 and 6:0 which is a time when children are learning and consolidating concepts of ordinal number. Before this period only a quarter of the sample passed and from 6 years onwards the proportion passing increased steadily with age from .86 to .97. Item 17 appears to be of equal difficulty on each Form.

From six years upwards the commonest mistake on Form A was to select the third animal in the line which is the second one from the other end. However, this mistake rarely occured on Form B where the trucks are identical; by far the commonest error on Form B ($p < .001$) was, at all ages, to point to the first truck in the array (i.e. the fourth from the sign). Similarly, children under six who failed Form A overwhelmingly selected the first animal ($p < .001$). This is also the largest animal.

It can be assumed that, apart from guessing, no child can pass item 17 without understanding the concept second though some could fail by identifying the second object from the wrong end of the array.

Item 18: corner

The Form B version is easier than that on Form A for this sample and for Boehm's. The depiction of buildings and streets from a high viewpoint does not seem to confuse. Form B depicts the concept more precisely: the building IS the corner of the street whereas the glass (Form A) is not touching the horizontal and vertical axes and so, geometrically, is not AT the corner even though it may be thus described in colloquial speech. It may be that some children have at least partial knowledge of corner but the extent of their grasp of the coordinates involved is not revealed by item 18. Piaget (1956) claims that the framework appropriate to comprehensive euclidean and projective systems is not completed until the age of nine or thereabouts. However, item 18 taps the child's knowledge of the colloquial use of corner.

Corner appears in all the curricula areas examined with the exception of Nursery Assessment.

Failure of item 18 implies lack of understanding of the colloquial usage and success indicates knowledge of it.

Item 19: several

Of the 50 BTBC items this is the weakest. The word several is rarely used in British infant schools and is therefore irrelevant to the purpose of the test as stated by Boehm (1971). Not all adults are entirely clear about the precise definition of several and it is used to denote various numbers of objects.

Moreover both forms of item 19 are badly designed. It is suggested that "several rabbits" (or spoons) may be confused with "seven". Some of the older children were observed muttering "seven" and looking confused and some replied "none of them". To the youngest children "seven" is generally understood as some number unknown i.e. "a lot" and therefore they could get item 19 right because they opt for the greater number of rabbits (or spoons) or because they hear the plural ending. Adopting the first-in-line strategy would achieve success on Form B. The Rasch analysis shows more of the low scorers succeeding than was expected and more of the high scorers failing than was expected.

Failure indicated incomprehension of the label several and in an unknown number of cases the inability to identify "more than one (usually more than three) but not very many" (Chambers, 1972) or "an indefinite number more than two, but not very many" (Webster, 1947).

Item 20: behind

Item 20 yields facility values which increase, and discrimination values which decrease, with each age level between 3:6 and 7:0. Some children found behind easier to locate on Form B (which is also noticeable in Boehm's sample). This is likely to be related to the necessity of shifting one's reference point from the front of the Form A picture (which is the side of the sofa) to the side of the picture (which is the front of the sofa); whereas the trailer (Form B), an object with an easily recognisable front, is between the boy to be identified as behind and the boy pulling the trailer.

Even though behind is conceived by H. Clark ( ) as the negative end of the front-back dimension it has been argued that placing an object behind another takes place at the positive end of the horizontal front-back dimension i.e. in front of the child (Kuczaj and Maratsos, 1975). It is probable that the object behind the other and therefore further towards the positive end of the dimension is conceptually easier and thus in such an instance behind should be regarded as a positive term (ibid.).

Clark (op. cit.) asserts that front is acquired earlier than back but some studies find that behind and back are comprehended earlier than front (Levine and Carey, 1982, Cox, 1979); the back axis is used sooner than the front one (Feagans, 1980). Durkin (1978, 1981) found that correct object-oriented responses for in front of but subjectively oriented responses for behind were common in infant school children. "It seems possible that this reflects in part the child's uncertainty with the terms, and in part the fact that in front of explicitly refers to the properties of the object, while behind does not." (Durkin, 1981 p.49). A complex disjunctive concept of the front-back orientation has been found to precede any knowledge of the words front and back (Levine and Carey, op. cit.). "Children apparently acquire the proper axis of reference as well as the notion of oppositeness of this axis prior to acquiring the individual meanings of the terms themselves" (Richards, 1979a p.33). Once one term is correctly interpreted so is the other, suggesting near simultaneous acquisition of both terms (Kuczaj and Maratsos, op. cit.)

By far the most frequent error on Form A ($p < .001$) was to identify the toy behind the sofa as the doll at the side of the sofa. Maybe some children, in this context, interpret behind as somewhere out of sight and indeed in some rooms a toy dropped between the sofa arm and the wall is colloquially described as behind. On the other hand, the doll is at the front of the picture and so the errors may indicate confusion of behind and front. It has been claimed that children at an intermediate state of lexical knowledge interpret front as if it means back (Levine and Carey, 1982). On Form B the commonest error was to choose the boy in front of the trailer but as the number of errors was small this trend is not statistically significant.

Success on item 20 implies understanding of the concept and its label but failure on Form A does not necessarily indicate absence of the concept behind; whereas failure on Form B would.

Behind features in all the curricula areas examined and is therefore correctly included in the BTBC.

Item 21: row

Though row is not often used in infant school work and was not observed in the nurseries the word row seems to have been learnt by at least 94% of children by the age of six. Item 21 discriminates well in the early years. It is of roughly equal difficulty on both Forms and therefore it is unlikely that children passed Form B simply by applying a first-in-line strategy. Success indicates comprehension and failure suggests lack of knowledge of row.

Item 22: different

As well as describing something which is unlike, different also has a usage which means separate, distinct or "another one of these objects". Therefore an item testing the concept different is likely to be context bound. In this case success may depend on paying close attention to all the words of the direction. A number of children (24%) who pass item 22 Form A fail Form B or vice versa. The correct answer on Form A, apparently, for this sample and for Boehm's (1971), is more obviously different from the rectangular assemblies of bricks than is the correct answer on Form B where the variation is in the height of the pile and the number of books. Most errors follow the pattern of choosing the first object in the array ($p < .05$ Form A; NS Form B).

Different features in all the areas of the curriculum and in teachers' speech with each age group.

Item 23: after

To achieve success on this item, theoretically requires linguistic understanding of a complex sentence, reversible thinking and the ability to master the memory load of a relatively lengthy sentence. Reversible thinking is decisive for correct before and after constructions (Trosborg, 1982). Sentences which preserve the actual order of events are easier than those which reverse the order; the most difficult of all are those of the construction 'B after A' as in item 23 (Weil and Stenning, 1978). Therefore a child who can distinguish after from before could fail this item because of the syntactical complexity of the direction.

Even though by far the commonest error was to point to the drawing which depicts the before situation ($p < .001$ Form A; $p < .01$ Form B) it cannot be assumed that these children, in other circumstances, would interpret after as before. Indeed there seems to be no stage where after is understood as meaning before (Coker, 1978); when four year olds understand either before or after they are aware that both terms refer to the same dimension in space (Cox, 1979).

It is probable that before and after are first acquired as prepositions and then as subordinating conjunctions (Coker, op. cit.) Before and after used as subordinating conjunctions present considerable difficulty for 5 and 6 year olds (Coker, 1978 ; French and Brown, 1977). The child may adopt either of two strategies: a syntactic strategy of attending to the main clause and not the subordinate (Coker, op. cit.; French and Brown, op. cit.) and a semantic strategy in which the order of mention of events is interpreted to respond to the actual order of occurrence (Coker, op. cit.; Clark, 1971). This semantic strategy is the more frequent when the child is cued to pay attention to both clauses; otherwise the syntactic strategy is more prevalent (Coker, op. cit.). As the children had already done 22 BTBC items some, at least, would be cued to attend to both clauses and therefore would be likely to apply the order-of-mention strategy to item 23. The use of either strategy does not imply partial knowledge or lack of knowledge of before and after. Contrary to Clark's (1971) assertion, when tasks are designed to control strategy effectiveness, there is no evidence that after is more difficult than before (Coker, op. cit.; French and Brown, op. cit.; Goodz, 1982; Trosborg, 1982).

There were fewer errors on Form B than on Form A. On Form B the picture showing the after situation is positioned last in the row. Children's comics place the last event in time at the end of a row of pictures. However, Boehm's (1971) results, which show more errors on Form B than on Form A, do not support this proposition. Even so, it is possible that some children chose the last and incorrect picture on Form A for the same reason. It is also conceivable that a few children may have been confused by the usage of after in the phrase 'look after' which was observed in teachers' discourse and in all of the reading schemes examined. Thus they might have chosen the picture showing the hairdresser 'looking after' the girl's hair. Others may have employed a first-in-line strategy by which they would have answered correctly on Form A and chosen the before situation of Form B.

Item 24: almost

In addition to almost item 24 also taps the concepts empty (Form A) and full (Form B). From 4 years onwards this is a very easy item on Form A with over 90% of each age group achieving success and virtually no discrimination above 3½ years. On Form B item 24 is difficult with a success rate of only 21%. Only 31% achieve the same score on both Forms. /Boehm (1971) records higher scores for Form B than Form A._7

A child can pass Form A without comprehending the word almost. As no bottle is completely empty the word almost may be disregarded and the correct answer arrived at from the cue empty or most empty. The majority of wrong responses ($p < .001$) are the first bottle in the row which is half empty. On Form B only 2 children selected the apparently empty basket whereas 107 chose the full basket. This may have been because, as with Form A, almost is not understood or is ignored, and the basket which is most full selected. Or, the compelling notion of fullness (Donaldson and McGarrigle, op. cit.) may have overridden other criteria and influenced judgments. It is likely that the children who did pass item 24, Form B understood the concept almost though some of these children may have used the first-in-line strategy. Failure on either Form indicates incomprehension of the word almost.

Almost appears infrequently in infant school curricula and seems to be absent from nursery curricula. It was only observed in teachers' discourse with seven year olds. Nearly and not quite are more common terms.

Not only is the word almost apparently unnecessary for success in the nursery and infant school and the design of Form A ill-conceived but the Form A drawings are of poor quality. The symbolic portrayal of water is not always understood by young children: 60% of the picture sample judged the second bottle, which is almost empty, to have no contents at all.

Item 25: half

Facility values on Form A range from .28 at 3½ to .88 at 7 years with the highest discrimination value (.40) at 6 years. The Form B version appears more difficult with an overall facility value of .39. Boehm (1971) also records higher scores for Form A. It seems therefore that the comprehension of half depends on context: young children are more familiar with a pie half gone than a box half black. Thus failure denotes a lack of understanding of half in certain contexts. Excepting those children who apply a first-in-line strategy a child would need to be able to judge half and know the label half in order to pass.

Errors on Form A tend towards a preference for the larger amount ($p < .05$) which is five sixths of the pie; both incorrect responses on Form B depict only a quarter of each box.

Half features throughout the nursery and infant school curriculum and is therefore a relevant concept.

Item 26: centre

This item apparently tests a child's knowledge of the concept label rather than of the concept itself. A number of older children were heard to mutter "That's the one in the middle"; many children asked "Do you mean the middle?" The word centre was only observed in teachers' discourse with seven and eight year olds. It does not feature in any of the mathematics workbooks examined nor in nursery curricula. It appears infrequently in one of the reading schemes and in Science 5/13.

Below 6 years there is no smooth progression in facility values. Discrimination values are generally low and are uneven across the age levels. The Rasch analysis reveals low scorers on the BTBC obtaining more passes than expected on item 26 and high scorers obtaining fewer passes than expected.

The commonest error on each Form was to select the object on the circumference of the circle. There may have been syntactical confusion i.e. the object with the circle through its centre or this may have been an expression of a preference for the topmost object.

The results suggest that item 26 is useful with seven year olds but has very limited uses with younger children.

Item 27: as many

Items 27 and 47 involve understanding of the numerical equivalence of sets without provoked correspondence (Piaget, 1952). At each age level item 27 presents less difficulty than item 47. Several factors may be responsible: firstly, the words as many may be more familiar than equal; secondly, the objects in item 27 seemed to be within the subitising range of the children who passed, with very few visibly counting them, whereas a much larger number of children counted the objects in item 47 (see Laxon, 1981); thirdly, item 27 needs a simpler comparative judgment: it provides a model to which one of three sets must be matched whereas item 46 requires the child to select two equal sets from three.

The proportion passing item 27 Form A increased sharply from .08 at 3:6 to .32, .57 and .73 at 4:0, 4:6 and 5:0 with 98% of 7 year olds passing. Discrimination is highest at 4:6 and 5:0. Both Forms are roughly equivalent in difficulty.

Many of the youngest children were unable to carry out the direction: they pointed to the box. Failure does not necessarily denote absence of the concept of numerical equivalence: a linguistically simpler task may well enable more of the younger children to pass. Failure does suggest an inability to comprehend the phrase has as many as x. Children are unlikely to pass without comprehending both this phrase and the concept it embodies.

As many rarely featured in the infant school work examined and not at all in the nurseries. Many was often observed; most frequently in the question "How many?" but also occasionally in phrases such as "too many", "not so many".

Item 28: side

Side presupposes knowledge of front and back for side refers to the axis which is not at the front-back axis of the object (H. Clark, 1973). Objects with no discernable front or back (e.g. the box on Form A) present a child with more difficulty than do objects with a discernable front-back axis (e.g. the car on Form B). Not surprisingly, therefore, the pass rates for the same children taking Forms A and B were respectively 55% and 98%. Boehm (1971) also records much higher scores for Form B. Moreover the two illustrations also differ in that one boy is drawn 4cm. away from the car whilst the equivalent circle is only 0.2cm. away from the box and therefore almost touching its corner. Hence to pass Form A a child must differentiate between corner and side whereas this is not a requirement for Form B.

Children who pass either Form can be said to understand at least one meaning of the word side but whether they interpret side as edge (or boundary) or as a position to the right or left of the front-back axis remains unclear. Passing item 28, on either Form, does not necessarily indicate acquisition of this last meaning and failing Form A does not preclude knowledge of it. It is extremely difficult to establish the correctness of a child's response to a featureless object: he may project an imaginary front on to it (Richards, 1979a). In the case of item 28 Form A, a child could perceive his view of the box as its side or as its top surface. If he sees it in the former position the logically correct answer would be the circle otherwise seen as inside the box.

As side features in all the curriculum areas examined and as discrimination and facility values are satisfactory the Form A version of item 28 (used with discretion) is of value but, where the tester wishes to identify the linguistically or conceptually immature amongst the youngest children, Form B would be a better choice.

Item 29: beginning

The different contents appear to cause variations in difficulty between Form A and Form B. Only 67% of the sub-sample made the same score on both Forms with 60% passing Form A and 79% passing Form B. Boehm (1971) also records higher scores for Form B. The boys climbing the stairs (Form B) are depicted in ascending order where the squirrels (Form A) are not though both the squirrel and the boy who are beginning to climb are situated at the beginning of the left to right array. Whereas the first boy is clearly shown with one foot on the bottom step the first squirrel is not so obviously about to climb. However, this did not present ambiguity to any of the 'picture' sub-sample (Table 73). The commonest error on Form A ($p < .001$) was to point to the third squirrel which is obviously climbing up the trunk. (The second squirrel was judged by 16 of the 30 children in the 'picture' sample not to be climbing.) The errors do not suggest that beginning is interpreted in terms of its polar opposite.

Neither beginning nor begins was observed often in nursery and infant school curricula: starting and starts are more common.

Success could be achieved by adopting a first-in-line strategy but failure indicates unfamiliarity with the word beginning or confusion with the concept.

Item 30: other

Other appears frequently in nursery and infant school curricula. Item 30 (Form A) becomes steadily less difficult with age, discriminating best below 6 years. Forms A and B are of approximately equal difficulty. It is unlikely that a child could pass item 30 without understanding the concept other or fail if he did understand.

Item 31: alike

This seems to be an unreliable item. Facility values are erratic across the age levels. Only 71% of the sub-sample made the same score on each Form; Form B had a higher pass rate (61%) than Form A (51%). Alike is rarely used in nursery and infant schools: the commonly used term is same. Some children may only attend to the second syllable and thus either choose the shapes which they like[1] or find features which are like other features such as the straight sides of the triangle, square and pentagon or the triangle and pentagon because, as one six year old said, "they've got those corners." For children who have recently explored mathematical shapes and extracted common features these are logical answers. The distinction between alike (or same) and resemblance (or likeness) is a fine one and beyond most young children except in highly specific contexts. With Form B, children are on more familiar ground: finding an identical pair of socks is a frequent if not daily task. Boehm (1971) reports slightly higher scores, in the kindergarten, for Form B.

As infant school children often sort geometric shapes and other objects into sets the difficulties encountered with item 31 are most likely to be

[1] Vicky, aged 4: This one's nice and this one's nicer.

semantic rather than conceptual. Perhaps some children reject the circles and the matching socks on the grounds that they are not two different things which are alike but are the same thing repeated. Thus a child with conceptual understanding of alike or same as may fail item 31 though she is unlikely to pass without at least some knowledge of both the concept and the label alike.

Item 32: not first or last

Item 32 apparently tests the child's knowledge of first, last, not first and not last. "Words containing Not (x) nearly always describe the other situation - the one that is not normal - and for that reason seem to be harder to acquire" (Clark, 1977 p.513). An added difficulty with the directions to item 32 is that young children frequently interpret only the first clause of a sentence (Townsend and Erb, 1975). In this case some children may have substituted but for or; others may only have 'heard' the final word last. Hence it is not surprising that roughly two thirds of the children who fail Form A select the last car. (Only just over half who fail select the last duck on Form B but the ducks and cars are travelling in opposite directions and some children may perceive the duck in the lead as the last duck.)

Item 32 in its present form cannot be justified as relevant to the test's declared purpose. The words first and last are encountered in the infant school and first is used by teachers and children in the nursery but the phrase not first or last was only encountered once (in one reading book) during the search of curricula materials and observation of teachers' discourse.

Item 32 requires an act of deductive reasoning. It may well be that the ability to reason from negative premises is closely related to a general language factor which could account for the higher than average discrimination (for the BTBC items) between 4½ and 6 years.

In addition to ignoring the final phrase some children with only partial understanding of the several concepts involved may be swayed into choosing the last car by their preference for the open sports car; whilst others may give the correct answer because of their preference for the second car (Table 71). Alternatively, some children may have thought that "Point to the car that is not the first or the last" gave them a choice of the second or last car.

Item 33: never

The drawings could be improved: until the tester named the biscuits (Form A) they were sometimes mistaken for flowers and buttons; the shoe could be more unisex.

For a number of children never seems to have the connotation of 'something forbidden' as in "Never play on the road" or "We never tear books". Thus, in the child's meaning, exceptions are accommodated.

Some children who understand the adult meaning of never may fail item 33 because of their egocentric interpretation of 'a child': for example, a six year old who did well on the rest of the test was later asked if he liked biscuits. "Yes" was the surprised answer. To the enquiry about why he had said that he never ate the biscuits he replied "I never eat biscuits with cherries on. I don't like them." Children with an incomplete understanding of never could still pass item 33.

Facility values, however, increase smoothly with age and discrimination values, though not high, are acceptable with the highest discrimination between 4½ and 6½ years. A fairly high proportion (.84) give the same response to both Forms though Form A is easier than Form B. Boehm (1971) also records higher scores for Form A. This may be because not all children wear watches and some may think that boys (or girls) do not wear shoes like the one depicted.

Never, used with the accurate adult meaning, features in the nursery and infant school but in addition to the colloquial usage noted above a local idiom was observed: mock horror or indignation expressed by "She never!" and denial: "I never".

Item 34: below

The results support Clark's (1977) findings that errors on below tasks are accounted for by a surface-based strategy i.e. when asked to place A below B 38% of Clark's subjects put A on the top surface and 22% put A on the same surface as B. In response to item 34 (Form A) 208 children put a finger on the table, 13 above it and 4 children touched the edge of the table. In response to Form B 61 children pointed to the bird perched on the back of the seat and 17 to the bird flying above the seat. These results conflict with Durkin's (1978) finding that in response to the direction "Put the brick below the table" no child put the brick above or on top of the table: all children in error placed it at the side.

Item 34, Form A discriminates best with the younger children. From six years very few children fail Form A which has a higher pass rate (87%) than Form B (65%). It is unlikely that young children encounter below in a context similar to that of Form B; everyday speech would refer to the bird under the seat; below may be more readily associated with a higher surface level such as a table. In addition, the bird supposedly below (or under) the seat might well be perceived as feeding on the ground in front of the seat. Below was not observed in teachers' discourse and does not appear in nursery curricula.

Boehm's (1971) findings show the reverse pattern of response with a higher success rate on Form B than on Form A. This may reflect some trans-atlantic difference in the usage of below when the referent is near ground level.

Success on either Form indicates understanding of below but failure on Form B does not necessarily mean that either the concept or its label is not understood.

Item 35: matches

Matches is not often used in nursery and infant schools; it was heard spoken only in one nursery. The concept is usually expressed by goes with or belongs.

The more familiar context of matching clothes (Form B) evidently facilitates greater success (.68) than does the context of Form A (.58). This is also reflected in Boehm's (1971) results. Only 61% of the sub-sample made the same score on both Forms. Thus a child who understands the concept matches in one (or more) contexts may fail item 35.

More than two thirds of the errors on Form A were the first ball and box which each have diagonal lines. Some children may have applied a first-in-line tactic whilst others may have learnt a restricted meaning of matches as 'bearing close resemblance' and so rejected the correct answer for being identical in pattern. Children who pass either Form of item 35 can be assumed to understand matches.

Item 36: always

There is a fundamental difference between the design of this item on each Form: a child may often, though not always, have a dog or a book whereas a bicycle, in most children's experience, never has a box as depicted in Form B nor a feather. Consequently the pass rate on Form B was 94% and on Form A only 53%. Boehm (1971) also reports similarly large differences between Form A and Form B. A child who equated always with often (or even sometimes) would be likely to pass Form B but his chances of passing Form A would be one in three.

As with item 33 (never) the child's egocentric interpretation may influence his response. Twice as many children erred on Form A by choosing the book rather than the dog. School children do have books and young owners of books or a dog could logically say that they always own or have them. Though item 36 (Form A) discriminates as well as many other BTBC items it may be testing a general verbal ability rather than specific knowledge of the concept always.

Always features in the infant curriculum but it was not observed in any nursery nor is it included in the Manual for Nursery Assessment.

Item 37: medium-/middle-sized

Neither medium-sized nor middle-sized was observed in teachers' discourse; nor do these terms appear in any of the curricula materials examined excepting mathematical vocabulary recommended to teachers. However, young children become familiar with the concept through many early experiences such as stories (The Three Bears for example), placing objects in order of size and through play with the sand and water equipment recommended by Parry and Archer (1975). As expected, the acquisition of this concept accelerates after 4½ years.

The two Forms appear to be of equal difficulty. As 23% of the sub-sample responded differently on Form A and Form B neither a pass nor a fail can be regarded as conclusive. A child could pass by applying a first-in-line strategy.

On both Forms the overwhelming majority of children ($p < .001$) who failed pointed to the second object. Some may have been influenced by a preference for the largest object; others may have attended only the the word middle and so selected the middle object in the array. This position is the one where the middle-sized object, animal or person is most often placed in children's book illustrations. So these responses to item 37 may be revealing a partial understanding of middle-sized (or medium-sized) in which order of physical placement is of more importance than sizing.

Item 38: right and item 44: left

Approximately 86% of the sample showed at least some understanding of the left-right orientation. Many children hesitated and some admitted "I always mix them up" and then obviously guessed. If children are uncertain as to which is left or right they have an even chance of guessing the correct answer. So children who pass either of these items may or may not be able to identify right and left. Even a quarter of such children who pass both items would be expected to do so purely by chance. If both alternate Forms are used then obviously the odds are reduced.

Additionally, in the absence of a firm knowledge of left and right, choices may differ from chance under the influences of handedness or the habit of looking first at the left-hand word or picture or the association of the word end with the right-hand side of the page. It is possible that some children 'heard' the direction to item 38 as "the box/apple right at the end of the line/shelf" despite the modified direction (App. b). This may be one reason why item 38 is passed by more children than is item 44 on both Forms. Alternatively, the semantic features hypothesis would predict that the positive term is learnt before the sinister. However, there is little difference between the frequencies of occurrence of these two terms in curricula materials; left was occasionally heard in teachers' speech whereas right was not.

Form B appears to be somewhat easier than Form A for both items. This also holds true for Boehm's results. The reasons are unclear. The greater number of objects in item 38, Form B may possibly aid success in that the eye can follow the sequence more easily or it may simply be that apples on a shelf are more meaningful than boxes and a line. The six year old who responded to item 44, Form A by asking "Do you mean my left or the birds' left?" may not have been unique in his conundrum.

Item 39: forward

Forms A and B are not strictly equivalent: whereas none of the chickens (Form B) appears to be bending backward one boy (Form A) is clearly doing so. The pass rate for Form B is 77% and a child could pass by attending to the word bending and ignoring forward. On Form A the pass rate is 56% and a child would be unlikely to pass without understanding forward. Boehm (1971) also reports a higher pass rate for Form B. Failure on either Form indicates a lack of understanding. Between 4 and 7 years item 39 (Form A) discriminates better than most other items.

Errors on Form A overwhelmingly (636:30) point to either forward being understood in terms of its opposite or being ignored and the response being made to bending.

Forward does not feature often in infant school work and, seemingly, not at all in the nursery.

Item 40: without any

This item discriminates well with the under-fives. The content validity study justifies the substitution of without any for zero. Item 40 is consistent across both Forms. Success indicates knowledge and failure suggests incomprehension of without any.

Item 41: above

This is another item in which there is a fundamental difference between Forms A and B. Form A shows an aeroplane below the cloud whereas there is no window below the door on Form B. Errors on Form A ($p < .02$) suggest confusion between above and below which cannot be revealed by Form B. Fewer errors were made on Form B by the sub-sample and by Boehm's (1971) sample.

The facility of the three items over (.94), above (.76) and below (.88) are not in the order predicted by E. Clark (1977) and H. Clark (1973). However, the order of acquisition suggested by the smaller sample who were tested with Form B does reflect their proposed acquisition order: on Form B the facility values for over, above and below are .93, .83 and .65. This apparent contradiction between the results from the two forms is thought to be caused by the differences in item content noted above and in the discussion of item 34.

Durkin (op. cit.) found almost twice as many errors for above as for below in a drawing task whilst his comprehension task exposed the children's uncertainty of the term above. He notes that above was rarely used by infant school children in an elicitation task (ibid.).

Item 41 would be a more searching test of children's understanding of above if each Form depicted an object on the referent as well as one at the at the same level and an object below it. At present, though children cannot pass merely by applying a surface strategy (Clark, 1980) the perceptual salience of the topmost object may enable some children to pass without a full knowledge of above. Failure, on the other hand, indicates incomprehension.

Item 42: every

Results of the content validity study show that every and combinations such as everything, are encountered frequently by young children.

Observations during testing suggest that errors in many cases were due to the syntactical complexity of the directions rather than incomprehension of every. Some children appeared to disregard the final three words in every bowl (circle). This is clearly reflected in the errors on Form B where twice as many children chose "the group that has a dot..." as chose the group with two dots ($p < .05$). As already noted this is a common strategy amongst young children.

So although discrimination and facility values suggest that item 42 is useful, especially from 3½ to 6 years, it may be more of a test of facility with syntactical complexity than a test of the comprehension of every. The fact that 23% of the sub-sample scored differently on each Form despite their similar facility values (.73 and .70) lends support to the view that item 42 is an unreliable test of a child's understanding of every. Children who do not comprehend every are unlikely to pass.

Item 43: separated

The inclusion of separated in the BTBC cannot be justified on the grounds that it is necessary to understand this word in order to achieve success in the first school. Separately was observed in classroom discourse only with the oldest children and occasionally it appears in mathematics books but neither separated nor separate was found in any area of the curriculum. Children probably first acquire understanding of together and apart or in pieces.

Facility and discrimination values suggest that item 43 is of most use with children aged six and seven. A child could pass by selecting the picture which looks different from the other two but, if his understanding of <u>separated</u> is secure, would be unlikely to fail. Some of the 28% of the sub-sample who gave different responses to each Form may have operated in this way (as may some of the children who passed both Forms); others appeared to guess. The commonest error on both Forms was to point to the first picture in the array ($p < .001$). In the case of Form B this could also be seen as <u>different</u> from the other two pictures.

<u>Item 44: left</u>

See item 38: <u>right</u>.

<u>Item 45: pair</u>

This is the most difficult of the 50 BTBC items with an overall facility value of .28. Both Forms are equivalent.

It seems that for most children <u>pair</u> is still an uncertain concept at 7 years of age despite familiarity with the word <u>pair</u> during the infant school years. Perhaps the word <u>pair</u> is associated with gloves and shoes but cannot yet be transferred easily to other objects.

On both Forms at least 90% of children who failed selected the first picture which contains three candles (or dolls). They may have been influenced by a preference for the larger amount, awareness of the plural ending or used a first-in-line strategy.

It is reasonable to assume that success on item 45 denotes a mature understanding of <u>pair</u> whilst failure suggests incomprehension or only a limited understanding.

Item 46: miss

The content validity study justifies the elimination of the word skip from item 46. Miss and its derivatives feature in reading and mathematics books and though miss was not observed in classroom discourse missing and missed out were heard. With hindsight it is suggested that the adapted instruction (Appendices 6 and 7) should be altered to "Miss out a box and point to the next box."

Both Forms are of equal difficulty with 90% of the sub-sample responding in the same way to each Form. Discrimination and facility increase around 5½ years and discrimination drops abruptly at seven. Below 5 years virtually no children were able to deal successfully with this item. Most of them pointed to the next box or ring. The errors of older children were generally of this nature also. Unlike the responses to items 32 and 42, where the last words of the direction are disregarded, with item 46 the final clause appears to be regarded as the important one. It is likely that during the preceding items children develop a 'response set' whereby the words "point to/show me" signal the important words of the sentence. Hence the first clause "Miss a box" is overlooked.

In addition to the concept miss item 46 also involves some knowledge of sequential ordering.

It is unlikely that a child could answer this item correctly without understanding miss. On the other hand, it is quite possible that many children who fail would succeed on some other task involving the same concept.

Item 47: equal

The content and construct of this item has been discussed in conjunction with item 27. As a correct solution entails the equivalence of sets without provoked correspondence it is not surprising that there is an upsurge around 6½ years in the number of children passing item 47 and the finding that only 75% of 7½ year olds pass is commensurate with Piaget's (1952) findings.

Item 47 is perhaps best viewed as a limited test of the numerical equivalence of sets rather than a test of the word equal which, in other contexts, might be understood by some of the children who fail this item. Before the word equal is taught in specific mathematical concepts the notions "same number" and "same amount" are generally familiar.

The reasons children gave for preferring one picture of lollipops to the others (Chapter 9, 13.07) support Hughes and Grieve (1980) in that a child will attempt to answer whatever question is asked however unanswerable or bizarre the question.

Item 48: in order

This is another item which is complicated by syntactic complexity. It is likely that many children who fail would be capable of successfully placing objects in order of size. Thus it is suggested that this item picks out those children who both understand the concept in order and can handle the syntax of the directions. The proportion of successes accelerates around 6½ years and discrimination is at its greatest at this age.

Item 49: third

Though Boehm (1971) labels the concept tested by item 49 as third the concept which it actually tests is third from x or N^{th} from x. Approximately half the 5 year olds passed item 17 (second) yet only 47% of the 6 year olds passed item 49. If item 49 solely tests third then these results would indicate that after acquiring second a further year is needed to acquire third which is nonsensical. It is likely that results similar to those already obtained with item 49 would be obtained if second or fourth were substituted for third.

By far the commonest error, on both Forms, is to point to either the first or the second child/house from the teacher/shop ($p < .001$). Pointing to the second child (or house) is either a further example of children disregarding the final words of the sentence or not understanding them and therefore including the teacher (or shop) in the ordinal sequence or else counting from the wrong end. No explanation has been found beyond the first-in-line strategy for almost as many children choosing the first child (or house).

Form A appears somewhat easier than Form B probably because of greater familiarity with children being first, second and third in a line. There is little difference between Boehm's (1971) results from each of the Forms.

Without understanding third from x a child would be unable to pass item 49. Though children with an understanding of third might fail they would not do so if they understood N^{th} from x.

Item 50: least

The results appear to show that least is acquired later than its unmarked opposite most (item 13). Whereas 47% of 3½ year olds and all 6½ and 7 year olds passed item 13 only 4% of 3½ year olds and 65% of 7 year olds passed item 50. However, as noted in the discussion of item 13, this is not empirical proof of an earlier understanding of most.

The majority of children who fail item 50 choose the most stars (Form A $p < .001$) and ice-creams (Form B $p < .01$). For some children this may be due to a preference for the larger amount or it may be that the meaning of most is overextended to include least. A child with a full understanding of least would be unlikely to fail item 50, though by applying a first-in-line strategy a child could pass.

Summary of content validity

Jensen's (1980) definition of content validity emphasises that test items should be representative of the specified universe of knowledge. Boehm (1971) defines this as (a) a group of concepts considered necessary for achievement in the first years at school and (b) which are "seldom if ever explicitly defined, or (are) defined in their simple forms but subsequently used in complex forms without adequate transitions" (ibid., p.3). Which concepts are used in complex form without adequate transitions is not clear nor can any evidence be found of this occurring in classroom practice.

Judged by these two criteria, on the basis of their appearance in teachers' discourse and in reading, mathematics and science books in first schools in England, the following 20 BTBC items are valid in content:

1. top	16. nearest	30. other
2. through	18. corner	33. never
4. next to	20. behind	36. always
5. inside	21. row	40. without any
7. middle	22. different	41. above
11. over	23. after	42. every
15. whole	28. side	

A further six items are accepted as valid by the first criterion but it is not taken for granted that they are understood by young children. Indeed much of the nursery and infant curriculum is planned around introducing and consolidating such concepts. This is seldom achieved by verbal definition alone: carefully structured play materials and activities together with the skills of an adult experienced in talking with young children, at the child's level of meaning, are required. Furthermore, it is questionable that any of the first group of concepts listed above, if misunderstood, would escape the attention of a skilled nursery or infant teacher. However, it is possible that adults inexperienced with very young children might make such assumptions.

The six items accepted as valid by the first criterion only are

13.	most	25.	half	49.	third from
17.	second	48.	in order	50.	least

If the concept labels were changed, a further six items would become valid by both criteria:

10.	around/round	29.	beginning/starting	34.	below/under
24.	almost/nearly	31.	alike/same/like	35.	matches/goes with/ belongs

A further seven items are deemed valid by the first criterion if the test is used with 6 and 7 year olds:

8.	few	14.	between	46.	miss (out)
9.	farthest	27.	as many		
10.	widest	37.	middle-sized		

The concepts embodied by these terms may be learnt earlier (in part at least) but not met in these forms in the reception class or nursery.

There is no evidence that understanding of the remaining eleven concept terms, in their present wording, is necessary for achievement in the first years at school (or at least in schools in England). These items are:

3.	away from	39.	forward
6.	some not many	43.	separated
19.	several	44.	left
26.	centre	45.	pair
32.	not first or last	47.	equal
38.	right		

Though children are familiar with the right-left dichotomy there is no evidence that correct identification of each term is required during early schooling. As noted earlier, the word <u>equal</u> is taught when required for

The six items accepted as valid by the first criterion only are

13.	most	25.	half	49.	third from
17.	second	48.	in order	50.	least

If the concept labels were changed, a further six items would become valid by both criteria:

10.	around/round	29.	beginning/starting	34.	below/under
24.	almost/nearly	31.	alike/same/like	35.	matches/goes with/ belongs

A further seven items are deemed valid by the first criterion if the test is used with 6 and 7 year olds:

8.	few	14.	between	46.	miss (out)
9.	farthest	27.	as many		
10.	widest	37.	middle-sized		

The concepts embodied by these terms may be learnt earlier (in part at least) but not met in these forms in the reception class or nursery.

There is no evidence that understanding of the remaining eleven concept terms, in their present wording, is necessary for achievement in the first years at school (or at least in schools in England). These items are:

3.	away from	39.	forward
6.	some not many	43.	separated
19.	several	44.	left
26.	centre	45.	pair
32.	not first or last	47.	equal
38.	right		

Though children are familiar with the right-left dichotomy there is no evidence that correct identification of each term is required during early schooling. As noted earlier, the word equal is taught when required for mathematical purposes. In particular, several is judged to be entirely without face validity.

In addition to antonyms of the BTBC concepts, such as bottom, outside, in front, which were used frequently by the teachers and are present in curriculum materials, the following concepts also appeared frequently:

apart/together	higher than	thick/thin
curved/straight	long/short	
edge	tall/short	

These seem more relevant to the stated purpose of the test than do some of the concept terms already included by Boehm.

Predictive validity

The results (Chapter 9, 9.00) suggest that the BTBC scores, scaled for age, can predict to some extent children's achievements, some nine to twelve months later, in mathematics, reading and spoken language (as assessed by teachers) with highly significant correlations of .48 (computation), .46 (mathematical concepts), .45 (speech) and .42 (reading).

These correlations between the BTBC and later assessments are compatible with correlations between the BTBC and subsequent achievement tests: SRA math. .56, SRA reading .43 (Piersel and McAndrews, op. cit.), median correlation of Stanford Achievement subtests (Estes et al., op. cit.), SAT spelling .54, SAT language .53, SAT arithmetic (computation) .58, SAT arithmetic (concepts) .65, (Steinbauer and Heller, op. cit.), Gates reading .56, (Busch, op. cit.).

The advantages and defects of teachers' assessments of the level of attainment of their pupils are well-documented (e.g. Schools Council, 1964). In early childhood it is likely that teachers' assessments are as reliable as standardised classroom tests. The reliability of the ratings made in mathematical purposes. In particular, several is judged to be entirely without face validity.

In addition to antonyms of the BTBC concepts, such as bottom, outside, in front, which were used frequently by the teachers and are present in curriculum materials, the following concepts also appeared frequently:

apart/together	higher than	thick/thin
curved/straight	long/short	
edge	tall/short	

These seem more relevant to the stated purpose of the test than do some of the concept terms already included by Boehm.

Predictive validity

The results (Chapter 9, 9.00) suggest that the BTBC scores, scaled for age, can predict to some extent children's achievements, some nine to twelve months later, in mathematics, reading and spoken language (as assessed by teachers) with highly significant correlations of .48 (computation), .46 (mathematical concepts), .45 (speech) and .42 (reading).

These correlations between the BTBC and later assessments are compatible with correlations between the BTBC and subsequent achievement tests: SRA math. .56, SRA reading .43 (Piersel and McAndrews, op. cit.), median correlation of Stanford Achievement subtests (Estes et al., op. cit.), SAT spelling .54, SAT language .53, SAT arithmetic (computation) .58, SAT arithmetic (concepts) .65, (Steinbauer and Heller, op. cit.), Gates reading .56, (Busch, op. cit.).

The advantages and defects of teachers' assessments of the level of attainment of their pupils are well-documented (e.g. Schools Council, 1964). In early childhood it is likely that teachers' assessments are as reliable as standardised classroom tests. The reliability of the ratings made in this study is not known but, judging by the interest and cooperation shown by the teachers, it is fairly safe to assume that the assessments were made conscientiously. There are no known factors related to the grading procedure that could have caused bias; though bias in favour, or against, individual children could not be controlled. The course grading (above average, average and below average) should have contributed to reliability: most variability in assessment is within the middle range of ability (ibid.) Even though it is impossible to obviate the halo effect entirely the steps taken probably reduced it. Not all children were given the same grade for each area of the curriculum.

Construct validity

A test is said to show construct validity if it predicts the behaviour that would be deduced from a theory of the trait being measured (Jensen, 1980). Cognitive theory predicts that with concept mastery a person would exhibit comprehension of the concept in a variety of contexts and would demonstrate mature understanding whereby the meaning of the concept is the same as society's meaning.

In order to establish perfect construct validity it would be necessary to demonstrate that each of the 50 BTBC items satisfies these criteria. It was beyond the scope of the present study to validate each concept in a variety of contexts. However, each concept was tested with alternative item content which is claimed by Boehm (1971) to be equivalent. Though the alternate forms of the BTBC yield roughly equivalent total scores the present findings are that 20 BTBC items show substantial differences between the facility values, on Form A and Form B, and another 6 items show some difference. Boehm's (1971) results also show similar differences between the two forms with many of these items.

Earlier in this chapter it was argued that these differences are caused by different contents and sometimes ambiguous drawings may have contributed to one form being found easier than the other. In a few cases different positioning of the correct answer may have enabled some children to pass on one form but not another if they employed a first-in-line strategy when faced with an item beyond their comprehension.

These variations in content and design may not be the only causes of difference between an item's facility values on Form A and Form B. Chance sample fluctuation may account for some of the differences though this is likely to have had minimal effect in cases where similar results are obtained by Boehm. Where an item is badly designed, on either or both forms, such defects can cause confusion and the child may be responding to a different question from the one actually asked by the tester. In such cases a different score on each form does not necessarily mean that the concept can only be understood in a limited range of contexts.

Where confusion is caused by linguistic complexity, the item, however satisfactory its facility and discrimination values, is not only testing a child's knowledge of the concept, it is also testing his ability to understand relatively complex syntax. Such an item may test either or both kinds of knowledge; additionally, the results may be confounded by the factor of simple forgetting in young children's responses (Kavanaugh, 1979; Weil and Stenning, 1978). The BTBC alone, even if both Forms A and B are used, does not distinguish between concept attainment and syntactical knowledge.

Neither does the BTBC distinguish between concept attainment and knowledge of the concept label. As argued in earlier chapters the two are neither synonomous nor often synchronous. At least five items (10, 24, 26, 31, 35) are probably tests of knowledge of the label rather than of the concept it represents.

A test's concurrent, predictive and construct validity is obviously reduced if children who lack the knowledge being measured make high scores and children who possess it make low scores. Arguments have been advanced as to which of these categories each item might be assigned. Excepting children who guess, the children who pass the following items can be assumed to understand the concept and its label, in this context at least, and children who fail these items can be assumed to fail because they lack a full understanding of the concept and /or do not comprehend the concept label:

Form A items

3.	away from	18.	corner	31.	alike[1]
5.	inside	21.	row	39.	forward
7.	middle	22.	different	40.	without any
9.	farthest	27.	as many	45.	pair
16.	nearest	30.	other	49.	third from

Form B items

3, 5, 9, 16, 18, 20 behind, 21, 22, 27, 28 side, 30, 31[1], 40, 45, 49.

Similarly, children who pass the following items can also be assumed to have conceptual understanding but some of these children may fail because of the design of the item; others will fail because they have not

[1] It is almost certain that, for most children, item 31 is a test of the word alike and not a test of the concept it represents. If it is regarded as a test of the concept then this item is placed in another category (NE) for both Form A and Form B.

yet acquired the concept or its label:

Form A items

14. between	28. corner	46. miss
17. second	34. below	47. equal
20. behind	35. matches[2]	48 in order
23. after	36. always	
26. centre[1]	42, evert	

Form B items

7 middle, 10 around, 14, 17, 23, 26[1], 34, 35[2], 42, 46, 47, 48.

Children who pass the following items may, or may not do so because they have reached conceptual understanding though the children who fail can be assumed to lack the necessary knowledge:

Form A items

1. top	19. several	41. above
8. few	24. almost	43. separated
11. over	25. half	44. left
12. widest	29. beginning	50. least
13. most	38. right	

Form B items

1, 4, 6, 8, 11, 12, 13, 19, 24, 29, 36 always, 38, 39 forward, 41, 43, 44, 50.

The remaining items neither reveal that a child understands the concept, in this or any other context, if he passes nor, necessarily, lacks conceptual

[1] Item 26 almost certainly tests knowledge of the word <u>centre</u> rather than the concept.

[2] Item 35 may well be a test of the word <u>matches</u>.

understanding if he fails. These items are:

Form A items

2.	through	10.	around	33.	never
4.	next to	15.	whole	37.	middle-sized
6.	some, not many	32.	not first or last		

Form B items

2. 15, 25 half, 32, 33, 37.

Figure 17a depicts the category into which each Form A item falls and Figure 17b the categorisation of Form B items. The first group of items identified above fall into the NW quarter, the second into the NE, the third into the SW and the fourth into the SE quarter.

<u>Form A</u>

passes reliable

N

3	5	7	9	14	17	20	
16	18	21	22	23	26	28	
27	30	31	39	34	35	36	
40	45	49		42	46	47	48

failures reliable W — E failures not reliable

1	8	11	2	4	6
12	13	19	10	15	32
24	25	29	33	37	
38	41	43			
44	50				

S

passes not reliable

<u>Fig. 17a</u> **<u>Form A items allocated to 4 categories</u>**.

Form B

passes reliable

N

3	5	9	16		7	10	14
18	20	21	22		17	23	26
27	28	30	31		34	35	42
40	45	49			46	47	48

failures reliable W ——— E failures not reliable

1	4	6	8		2	15	25
11	12	13	19		32	33	37
24	29	36	38				
39	41	43	44				
50							

S

passes not reliable

Fib. 17b Form B items allocated to 4 categories.

Even though the reasons advanced for the categorisation of each item are largely based on speculation, which may be disproved by further research, it is clear that not every item is a good discriminator. Where doubt exists about children's interpretation of language it is a sound axiom to assume that some children are not making an adult interpretation. This being so, it is a well-nigh impossible task to construct a picture test where each item measuring understanding of a different concept will be fool-proof. Either, additional forms are required, or the number of concepts should be reduced to allow more than one item per concept.

Complete acquisition of a dimensional concept cannot be presumed until the child has contrasted it with at least one other concept from the dimension (Clark, 1980) and been retested in different contexts. Both linguistic and extra-linguistic contexts influence a child's responses (Mills and Funnell, 1983; Light, 1983; Olson and Torrance, 1983; Robinson, 1983; Donaldson, 1982, 1978; Gelman, 1982, 1979; Hargreaves et al., 1982; Hughes and Grieve, 1980; Harris et al., 1978; Hoogenraad et al., 1978; Siegel, 1978; Weil and Stenning, 1978; Townsend and Erb, 1975; Campbell and Wales, 1970)

If the separate items are not all reliable tests of the concept they purport to measure what does the total BTBC score reveal about a child's present or future achievements?

Attempts to identify the factors contributing to the BTBC scores by use of principal components analysis were unfruitful. Neither Varimax nor Promax rotations produced consistent factor loadings and no meaningful interpretation could be made of the nature of the large number of factors. Other factor analytic studies have produced claims that the BTBC is a single, general factor related to the acquisition of basic verbal concepts (Piersel and Reynolds, 1981), that its main components are perceptual-motor and verbal abilities (Brown, 1975), that it functions as a general estimate of cognitive development and receptive language abilities (Steinert, 1978).

The present analysis of the BTBC items suggests that the following components contribute to the total BTBC score:

a) mature conceptual understanding,

b) partial conceptual understanding,

c) lexical knowledge.

d) facility with complex syntactical constructions,

e) auditory receptive and association skills[1],

f) visual receptive and association skills[2],

g) the abilities to quantify, compare, order and seriate,

h) reversible thinking,

All of which could be considered to be components of general cognitive development.

Some support for these conclusions is offered by an earlier principal components analysis of measures of four year olds' abilities and attributes (Smith, 1977). This identified a first order factor (Factor II) on which the BTBC, the ITPA mean scaled scores and previous Binet non-verbal scores loaded .84, .74 and .60 respectively. (The only other variables which loaded more than .40 on this factor were ITPA Auditory Association gains (.60) and previous nursery attendance (.60).) A second order factor showed a strong language component with the ITPA MSS and the BTBC loading .77 and .70 and previous scores for Mean Length of Utterance, language complexity and Binet non-verbal items loading .68, .53 and .51.

Further support is given by concurrent and predictive correlational evidence which shows high correlations of scores from the BTBC and the ITPA (ibid.), the Preschool Inventory and the Quick Test IQ (Ernhart et al., 1977) and arithmetical concepts (Steinbauer and Heller, 1978); moderately strong relationships have been found between BTBC scores and counting and computation skills (Piersel and McAndrews, 1982; Estes et al., 1976;

[1] Auditory association is the organising process of manipulating linguistic symbols (presented vocally) in a meaningful way (Kirk, McCarthy and Kirk, 1968). In other words, it is the ability to relate concepts presented orally (ibid.).

[2] Visual association embodies the ability to relate concepts presented visually (ibid.).

Steinbauer and Heller, op. cit.), the PPVT (Ernhart, op. cit.), Reading (Busch, 1980; Piersel, op. cit.), Language and Spelling (Steinbauer, op. cit.) and teachers' assessments of children's achievements in mathematics, reading and language made for the present investigation.

Additional to the components suggested above are the hidden factors which allow some children to perform well in a test and place others at a disadvantage (Anastasi, 1981; Wohlwill and Heft, 1977; Tizard, 1974; Zigler and Butterfield, 1968).

In sum, the BTBC may be described as a multi-dimensional test which may be considered to be as much a measure of young children's general cognitive development as it is of their understanding of the Boehm concepts.

Semantic Features Acquisition theory and non-linguistic practices.

A recurrent problem with comprehension tasks is that young children's performance is confounded by their response biases to linguistic and non-linguistic contexts. The role of such strategies in the acquisition of relational terms has been hypothesised by Eve Clark (1973, 1977).

The three general developmental principles of Clark's theory have not all received consistent support from the many studies it has generated. The most comprehensive of these principles which assumes that children learn the meanings of words in a top-to-bottom order (i.e. from the broader semantic features to the more specific features) has received consistent support for terms with a spatial reference (Richards, 1979 a, b; Carey, 1978: Brewer and Stone, 1975; Eilers et al., 1974; Campbell and Wales, 1970; Donaldson and Wales, 1970).

Clark's 'top-to-bottom' order of acquisition of pairs of terms that appear in the BTBC are top - bottom, in front - in back (behind), in (side) - out, first - last, over - under, above - below, before - after. The facility values obtained suggest that top (.97), inside (.94) and over (.94) are acquired before above - below (.76, .88) which in turn seem to be learnt earlier than after (.43) but behind (.80) does not follow this order. However, there could be a cross-cultural difference with 'in back of' and 'behind' (Durkin, 1981, 1978).

A second principle of Clark's theory is the asymmetric acquisition of semantically contrasting terms whereby the unmarked (or positive) term of a pair is learnt before its marked (or negative) opposite. This hypothesis is derived from the top-to-bottom principle: the marked term is regarded as semantically more complex than the unmarked one. Although there is evidence which suggests that some unmarked terms are learnt before their opposites (e.g. Clark, 1977 ; Donaldson and Balfour, 1968; Townsend, 1976; Hudson et al., 1982) there is also evidence of the marked term being learnt first (Levine and Carey, 1982) and of simultaneous acquisition (Kuczaj and Maratsos, 1975). Eilers (1970) reports that while some children clearly focus on unmarked adjectives others focus on the marked terms. It is possible that some children learn the positive term first whereas others learn the negative term first (Coker, 1978).

The facility values obtained suggest that these unmarked BTBC terms are learnt earlier than their marked opposites ($p < .001$): next to (.94) before away from (.81), nearest (.98) before farthest (.85), most (.87) before least (.35), whole (.75) before half (.62), right (.56) before left (.48). However, it is impossible to generalise from these BTBC results as other variables, such as item content, may be confounding. For example, on Form A, above (positive) was found to be more difficult than below (negative) whereas on Form B the reverse order of difficulty was observed.

The third principle of Clark's theory is the semantic over-extension of the meaning of the unmarked term. As noted earlier in this chapter the evidence that appears to support this is not conclusive (Hudson et al., op. cit.).

Clark suggests that "children start off by relying on their non-linguistic strategies to make a preliminary mapping from concept to word, and depending on how good the fit is, they will find the ultimate mapping more or less complex. The least complex mapping will be for those words where there is direct overlap of nonlinguistic concept or preference and adult meaning" (Clark, 1980 p.337).

Children's choices of wrong responses to the BTBC items were examined. It was hypothesised that non-linguistic preferences predicted by Clark's theory would predominate. Evidence was found for the following preferences:

a) A preference for the greater amount was observed in the wrong responses to item 25A ($p < .05$) and item 50 ($p < .001$). Errors on items 9A, 17A, 37 and 45 ($p < .001$) also show this dominance but the greater amount in these cases is also either the first object in the array (17A and 45) or the semantically pertinent position (37) or the nearest object to the referent (9). The greater amount also dominates the errors in two other items on Form A ($p < .001$) but not Form B (items 6, 8).

The greater amount did not appear to dominate the errors in item 24A but the poor drawing was misinterpreted by 40% of the sample who were questioned about it.

b) The perceptual salience of the vertical over the non-vertical (H. Clark, 1973; Clark and Clark, 1977) dominates the errors in Form A items 34 ($p < .001$) and 41 ($p < .02$). (The test was not applicable to 34B and 41B.)

c) The upper end of the vertical axis (ibid.) did not, however, dominate the errors in item 29 though it appears to have affected responses to Form A item 26.

d) The perceptual salience of nearness over farness dominates the errors in item 9 on both Forms ($p < .001$). On Form A the nearest object is also the largest but on Form B there are virtually no variations in size.

e) A non-linguistic strategy of choosing the first object in the line appears to have been employed with items 1A, 7, 17, 24A, 31, 35A, 38 and 43 ($p < .001$), 22A ($p < .01$) and item 31 at ages three and four ($p < .001$). (Errors on Form B items 1, 5, 13 and 22 are in the same direction but do not reach statistical significance.) On the other hand, not only do the errors on some items not show first-in-line dominance but the frequency of first-in-line errors is significantly lower than that of other errors on Form A items 14, 32, 36 and 39 and Form B items 10, 23 and 42. Items 15A, 19A and 45 ($p < .001$) also show first-in-line errors predominating but in the case of item 45 there may be other more compelling reasons for the choice (one linguistic and the other a preference for the larger amount). If a first-in-line strategy had operated exclusively on items 15 and 19 it should have produced more success on Form B where the correct answer is the first one.

In sum, the findings of this study in relation to linguistic and non-linguistic preferences and strategies are equivocal. It is likely that, when faced by difficult linguistic tasks, young children are swayed sometimes by one preference or strategy and sometimes by another according to context and for some children one particular preference or strategy may dominate for a period of time. If the data were reworked it might be possible to detect groups of children who consistently employ one, or more, strategies when faced with the BTBC items. However, the results would probably still be equivocal as other variables intrude and confound.

More importantly, in the absence of evidence that such strategies are not employed, it behoves whoever seeks to interpret the results of the BTBC (or any other comprehension test) to be wary of assuming that children's responses are made entirely on a linguistic basis. This is also an important consideration for test constructors.

e) A non-linguistic strategy of choosing the first object in the line appears to have been employed with items 1A, 7, 17, 24A, 31, 35A, 38 and 43 ($p < .001$), 22A ($p < .01$) and item 31 at ages three and four ($p < .001$). (Errors on Form B items 1, 5, 13 and 22 are in the same direction but do not reach statistical significance.) On the other hand, not only do the errors on some items not show first-in-line dominance but the frequency of first-in-line errors is significantly lower than that of other errors on Form A items 14, 32, 36 and 39 and Form B items 10, 23 and 42. Items 15A, 19A and 45 ($p < .001$) also show first-in-line errors predominating but in the case of item 45 there may be other more compelling reasons for the choice (one linguistic and the other a preference for the larger amount). If a first-in-line strategy had operated exclusively on items 15 and 19 it should have produced more success on Form B where the correct answer is the first one.

In sum, the findings of this study in relation to linguistic and non-linguistic preferences and strategies are equivocal. It is likely that, when faced by difficult linguistic tasks, young children are swayed sometimes by one preference or strategy and sometimes by another according to context and for some children one particular preference or strategy may dominate for a period of time. If the data were reworked it might be possible to detect groups of children who consistently employ one, or more, strategies when faced with the BTBC items. However, the results would probably still be equivocal as other variables intrude and confound.

More importantly, in the absence of evidence that such strategies are not employed, it behoves whoever seeks to interpret the results of the BTBC (or any other comprehension test) to be wary of assuming that children's responses are made entirely on a linguistic basis. This is also an important consideration for test constructors.

Recent work on semantic development emphasises the context dependency of young children (Harris, Macrae and Basset, 1978; Hoogenraad, Grieve, Baldwin and Campbell, 1978). Non-linguistic information interacts with lexical and syntactic knowledge (Donaldson and McGarrigle, 1974). Syntactic strategies appear to be independent of the meanings of terms (Coker, 1977). In the comprehension of adult speech, younger children rely more upon non-linguistic information than do older children and the range of this information is wider than was once supposed (Wilcox and Palermo, 1982). Thus, the construction of a test, of young children's comprehension of concept terms, which will prove reliable is a difficult exercise.

CHAPTER 12

General Conclusions

The main objective of the research has been achieved. The Boehm Test of Basic Concepts (with the directions slightly adapted) has been standardised on children in England. The total BTBC scores have been scaled by age to produce norms for each six month age level from 3½ to 7½ years which are applicable to both Form A and Form B. The Form A raw scores have a mean reliability estimate of .81 (KR_{20}) and a mean standard error of measurement of .27. No overall sex differences were found in the BTBC scores; nor were any significant differences found between the scores of children tested early in the school year and those tested during the second half of their school year.

Two short forms of the BTBC Form A (Tests C and D), were compiled for children between 3½ and 5 years. Correlation coefficients between the scores from each of these short tests and the scores from the whole test are .94 for Test C and .89 for Test D but these are bound to be inflated since Tests C and D are subsets of the whole test and share common items.

Boehm (1971) claims that the BTBC measures children's mastery of concepts considered necessary for achievement in the first years at school. As argued in the last chapter, the present findings suggest that, either on Form A or Form B or on both Forms, certain BTBC items are not reliable indicators of concept mastery. If a child passes, say, twenty items it is far from certain that he understands all twenty concepts or, conversely, if he fails twenty items that he has no understanding of these twenty concepts. Success depends on more than acquisition of the concept represented by the concept label. It also requires lexical knowledge and the ability to interpret syntax. Furthermore, inadequate item design may confuse the issue.

Cognitive theorists, Piagetian or otherwise, agree that concept acquisition is a gradual process. Models of concept learning proposed by Vygotsky (1962) and by CLD theory (Klausmeier, 1979b), for example, suggest that concepts are learnt at successively higher levels. Piagetian theory emphasises that concepts do not develop in closed systems but are in constant interaction with each other (Inhelder, 1974). Through assimilation and accommodation fresh cognitive structures are created at increasingly more complex levels. (See Chapter 3 A, p.14) So conceptual understanding is likely to be context bound to some degree until complete mastery has been achieved. Conceptual knowledge that is still incomplete may yet be adequate for understanding the directions encountered during the first years at school. A BTBC item should be viewed as an indicator of a child's understanding of the concept term <u>in the context of the item as presented</u> and not as an indication of concept mastery. If both the content and structure of what is said are regarded as dependent variables the context is the independent variable.

Observation of teachers and examination of curriculum materials suggest that mastery of all 50 concept terms is not essential for success in British first schools. However, most children are likely to encounter them before their eighth birthday.

The BTBC scaled score is considered to be a measure which estimates general cognitive development. Correlational evidence, from this and other studies, supports this view. As suggested in Chapter 11, understanding of basic concepts, whether partial or complete, is considered to be among the components of this general estimate of cognitive functioning.

Children's responses to BTBC items were examined in relation to Clark's Semantic Features theory. No conclusive evidence was offered. Even though the observed responses appear to be generally in accord with Clark's findings no contribution can be made to the semantic features debate in the absence of alternative experiments which are repeated with the same subjects with the variables controlled. Even if such experimental evidence were forthcoming its interpretation would not be straightforward. We can only observe behaviour and our interpretations of children's responses to tasks or questions may go astray.

Nevertheless, the evidence so far, both from the literature (see Chapters 3 B and 4) and from this study reveals that children's performance on linguistic tasks (including the BTBC) is influenced by non-linguistic information. Thus a child's performance on the BTBC needs to be interpreted with due caution and in conjunction with other evidence.

As predicted, performance on the BTBC is related to social class but there is considerable overlap of scores across SES groups. Thus socio-economic status cannot efficiently predict a child's performance.

Differences were found between the performance of children who attend schools in different types of location. Children of the same SES group tend to get higher BTBC scores if they are in rural areas which are not markedly socially disadvantaged whilst the children in urban rehousing estates tend to perform less well than their counterparts in the inner city or in other urban areas. The relative success of the inner-city children is, to a large degree, attributed to their schools. (This does not imply that the teaching in other schools is defective.) These results together with observation of the inner-city schools which were sampled suggest that it is time to update the picture of EPA schools which was presented by the Plowden report and the EPA research which followed it. (See Chapters 6 and 10)

The relatively poor performance of children in rehousing estates outside the inner core of the city or conurbation has implications for social, housing and educational policies. However, not all cities are alike and it would be rash to generalise.

Finally, what are the uses to which the BTBC may be put in Great Britain? The present adaptation and standardisation applies to children in England; it is not known whether populations in Ireland, Scotland and Wales would perform differently but the English scales can be utilised throughout the UK with more confidence than can Boehm's. Comparison of British scores with the US norms is hazardous for not only is Boehm's sample admitted to be non-representative of the US but the US BTBC norms are presented by school grade level and are not age-related in the way they are presented here.

Boehm (1971) declares that the BTBC can be used to identify children whose overall level of concept mastery is low and to identify concepts which are unfamiliar to large numbers of children in a class. This research suggests that the BTBC can be used to identify

a) nursery and first school children whose general cognitive development is below the mean for their age group,

b) children between 3½ and 6 years[1] whose general cognitive development is average or above,

c) concepts which may be unfamiliar to individuals or groups of children.

Individual administration of the test is advised. One of the shorter forms (Tests C and D) is recommended for the youngest children.

Educationists, psychologists and researchers will each decide on the purpose for which the test is required but interpretation of any results should be made in the light of the caveats presented earlier. In particular, the identification of individual concepts needs extreme caution and verification by other means is essential.

[1] A ceiling effect operates after this age. (See Fig. 3)

Few nursery and infant teachers will require this last warning. Indeed the majority of teachers in British nurseries and infant schools are well aware of incomplete conceptual understanding in their pupils. The activities they provide and the convers ations they engage in with the children are structured accordingly. Developing conceptual understanding is at the core of the curriculum.

<u>Summary</u>

1. The BTBC has been standardised for use with children in England between the ages of 3½ and 7½.

2. Norms for each 6 month age level in the form of T scores are presented as well as the proportion of each age group passing each of the 50 items.

3. Two shortened forms of the BTBC have been standardised for use with children between 3½ and 5 years.

4. The total BTBC score appears to provide an estimate of general cognitive development as well as providing a general guide to concept acquisition.

5. The reliability of each of the 50 items on each of the two alternative Forms (A and B) is variable. Relatively few of the test items can be regarded as reliable indicators of <u>both</u> acquisition and non-acquisition of the particular concept being tested.

6. Both non-linguistic and linguistic information may produce misleading results. Where such possibilities exist the test items are identified.

7. The order of acquisition of the BTBC concept terms appears to be generally in accord with the order predicted by Clark (op. cit.).

8. Performance on the BTBC is related to SES group membership.

9. Relationships are reported between domiciliary areas and performance on the BTBC. Children in non-disadvantaged rural schools tend to achieve higher scores than do children from the same social class who go to other schools. Children in urban rehousing estates tend to make low scores. Children in the inner-city schools in this sample performed better than was expected from earlier studies.

10. The BTBC is quickly and easily administered and children enjoy it. Now that British norms are provided, and the directions modified, it can be used in the UK as a research instrument and as a screening device. It can now also be used with preschool children.

BIBLIOGRAPHY[1]

ANASTASI, A. (1976) Psychological Testing. Fourth Edition. New York: Macmillan Pub. Co., Inc.

ANASTASI, A. (1981) Coaching, Test Sophistication, and Developed Abilities. Am. Psychol., 36, 10, 1086-93.

ANASTASI, A. (1982) Psychological Testing. Fifth Edition. New York: Macmillan Pub. Co., Inc.

ANDERSON, J.R. (1982) Acquisition of Cognitive Skill. Psychol Rev., 89, 4, 369-406.

ANGLIN, J.M. (1977) Word, Object, and Conceptual Development. New York: W.W. Norton & Co., Inc.

ANTHONY, W.S. (1977) Activity in the Learning of Piagetian Operational Thinking. Br. J. educ. Psychol., 47, 18-23.

AULT, R.L., CROMER, C.C. and MITCHELL, C. (1977) The Boehm Test of Basic Concepts: A three-dimensional version. J. educ. Res., 70, 4, 186-188.

[1] Periodical title abbreviations, mostly, follow the general principles recommended by the British Psychological Society; otherwise, American abbreviations have been taken from Periodical Title Abbreviations: By Abbreviation 4th Edition. Vol. 1 (ed.) L.G. Alkire, Jr. (1983) Gale Research Co: Detroit, Michigan.

BAGLIN, R.F. (1981) Does "Nationally" Normed Really Mean Nationally? J. educ. Measurement, 18, 2, 97-107.

BANDURA, A. (1977) Social Learning Theory. Engelwood Cliffs, N.J.: Prentice-Hall.

BANDURA, A., ROSS, D. and ROSS, S.A. (1967) A comparative test of the status envy, social power and secondary reinforcement theories. In Lazarus, R.S. and Opton, E.M. (eds.) Personality. Harmondsworth: Penguin.

BATE, M. and SMITH, M. (1978) Assessment in Nursery Education. Windsor: NFER Pub. Co. Ltd./DES.

BECKER, W.C. (1977) Teaching Reading and Language to the Disadvantaged - What we Have Learned from Field Research. Harv. Educ. Rev., 47, 4, 518-543.

BECKER, W.C. (1978) The National Evaluation of Follow Through. Behaviour-Theory-Based Programs Come Out on Top. Educ. Urb. Soc., 10, 4, 431-459.

BEILIN, H. (1980) Piaget's Theory: Refinement, Revision, or Rejection? In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking. New York: Academic Press.

BELL, C. and NEWBY, H. (1978) Community, Communion, Class and Community: The Social Sciences of the New Urban Politics. In Herbert D.T. and Johnson, R.J. (eds.) Social Areas in Cities, Vol. II. Chichester: John Wiley & Sons.

BELLER, E.K. (1973) Research in Organised Programs of Early Education. In Travers, R.M.W. (ed.) Second Handbook of Research on Teaching. Chicago: Rand McNally & Co.

BENTLER, P.M. (1980) The Study of Cognitive Development through Causal Modeling with Qualitative Data. In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking. London: Academic Press, Inc. (London) Ltd.

BEREITER, C. and ENGELMANN, S. (1966) Teaching Disadvantaged Children in the Preschool. New Jersey: Prentice-Hall Inc.

BERNSTEIN, B. (1970) Education Cannot Compensate for Society. New Society, 15, 344-347.

BERNSTEIN, B. (1971) Class, Codes and Control, Vol. I. London: Routledge & Kegan Paul.

BIERSWISCH, M. (1970) Semantics. In Lyons, J. (ed.). New Horizons in Linguistics. Harmondsworth: Penguin.

BIGGS, E. (1971) Mathematics for Younger Children. London: Macmillan.

BLANK, M. (1968) Experimental Approaches to Concept Development in Young Children. In Lunzer, E.A. and Morris, J.F. (eds.) Development in Human Learning. London: Staples Press.

BLEWITT, P (1983) Dog versus Collie: Vocabulary in Speech to Young Children. Devel. Psychol., 19, 4, 602-609.

BLOCK, J. (1982) Assimilation, Accommodation, and the Dynamics of Personality Development. Child Dev., 53, 281-295.

BODEN, M.A. (1982) Is equilibration important? - A view from artificial intelligence. Br. J. Psychol., 73, 165-173.

BOEHM, A.E. (1966) The Development of Comparative Concepts in Primary School Children. Unpublished Ph.D., 67-5767: Columbia University.

BOEHM, A.E. (1971) Boehm Test of Basic Concepts. Forms A and B. Manual 1971 Edition. New York: The Psychological Corporation.

BOEHM, A.E. (1980) Summary of Work with the Boehm Test of Basic Concepts. Monograph (in process).

BOEHM, A.E., KAPLAN, C.H. and PREDDY, D. (1980) How Important are Basic Concepts to Instruction: Validation of the Boehm Test of Basic Concepts. Teachers College, Columbia University (in process).

BRAINERD, C.J. (1979) Concept Learning and Development Stage. In Klausmeier, H.J. and Associates. Cognitive Learning and Development. Cambridge, Mass.: Ballinger Pub. Co.

BRAINERD, C.J. (1983) Modifiability of Cognitive Development. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

BRANDIS, W. and BERNSTEIN, B. (1974) Selection and Control: teachers' ratings of children in the infant school. London: Routledge & Kegan Paul.

BRANDIS, W. and HENDERSON, D. (1970) Social Class, Language and Communication. London: Routledge & Kegan Paul.

BREARLEY, M. and HITCHFIELD, E. (1966) A Teachers' Guide to Reading Piaget. London: Routledge & Kegan Paul.

BREARLEY, M. et al. (1969) Fundamentals in the First School. Oxford: Basil Blackwell.

BREWER, W. and STONE, J.B. (1975) Acquitition of spatial antonym pairs. J. exp. Psychol., 19, 299-307.

BRESLOW, L. (1981) Reevaluation of the Literature on the Development of Transitive Inferences. Psychol. Bull., 89, 2, 325-351.

BRODY, G.H. and ZIMMERMAN, B.J. (1978) Concept Formation. A Social Process. Educ. Urb. Soc., 10, 4, 459-76.

BRONFENBRENNER, U. (1977) Ecological factors in human development in retrospect and prospect. In McGurk, H. (ed.) Ecological Factors in Human Development. Amsterdam: North-Holland Pub. Co.

BROPHY, J. (1970) Mothers as teachers of their own preschool children: the influence of SES and task structure on teaching specificity. Child Dev., 41, 70-94.

BROWN, D.R. (1975) Validation of the Boehm Test of Basic Concepts. Unpublished Ph.D.: University of Wisconsin-Madison.

BROWN, G. and DESFORGES, C.W. (1977) Piagetian Psychology and Education. Br. J. educ. Psychol., 47, 7-17.

BROWN, R. (1973) A First Language. London: Allen & Unwin.

BRUNER, J.S. (1966) On Cognitive Growth. In Bruner, J.S., Olver, R.R. and Greenfield, P.M. Studies in Cognitive Growth. New York: John Wiley & Sons, Inc.

BRUNER, J. (1977) The Course of Cognitive Growth. In Child, D. (ed.) Readings in Psychology for the Teacher. London: Holt, Rinehart & Wilson.

BRYANT, P.E. (1974) Perception and Understanding in Young Children. Andover: Methuen.

BRYANT, P. (1982) The role of conflict and agreement between intellectual strategies in children's ideas about measurements. Br. J. Psychol., 73, 243-251.

BRYCE, T.G.K. (1981) Rasch-Fitting. Br. educ. Res. J., 7, 137-153.

BUSCH, R.F. (1980) Predicting First Grade Reading Achievement. Learning Disabil. Q., 3, 38-48.

CAMPBELL, R. and WALES, R. (1970) The Study of Language Acquisition In Lyons, J. (ed.) New Horizons in Linguistics. Harmondsworth: Penguin.

CAREY, S. (1978) Less May Never Mean More. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

CARR, D.B. (1979) The development of young children's capacity to judge anomalous sentences. J. Child Lang., 6, 2, 227-241.

CATON, H.R. (1977) The Development and Evaluation of a Tactile Anolog to the Boehm Test of Basic Concepts Form A. J. Visual Impairment and Blindness, 71, 382-386.

CENTRE FOR EDUCATIONAL DISADVANTAGE (1980) Educational Disadvantage in Rural Areas. Manchester: Centre for Educational Disadvantage.

CHAMBERS - MACDONALD, A.M. (ed.) (1972) Chambers Twentieth Century Dictionary. New Edition. Edinburgh: W. & R. Chambers Ltd.

CHAZAN, M., LAING, A.F. and JACKSON, S. (1971) Just Before School. Oxford: Basil Blackwell.

CHOMSKY, N. (1981) Knowledge of Language: its elements and origins. In Longuer-Higgins, H.C., Lyons, J. and Broadbent, S. (eds.) The Psychological Mechanisms of Language. London: The Royal Society and the British Academy.

CLARK, E.V. (1971) On the acquisition of the meaning of before and after. J. Verb. Learn. Verb. Behavr., 10, 266-275.

CLARK, E.V. (1973) What's in a Word? In Moore, T.E. (ed.) Cognitive Development and the Acquisition of Language. London: Academic Press.

CLARK, E.V. (1977) Strategies and the Mapping Problem in First Language Acquisition. In Machamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

CLARK, E.V. (1980) Here's the top: non-linguistic strategies in the acquisition of orientational terms. Child Dev., 51, 329-338.

CLARK, H.H. (1973) Space, time, semantics and the child. In Moore, T. (ed.) Cognitive Development and the Acquisition of Language. New York: Academic Press.

CLARK, H.H. and CLARK, E.V. (1977) Psychology and Language. New York: Harcourt Brace Jovanovich, Inc.

CLARKE, J., HALL, S., JEFFERSON, T. and ROBERTS, B. (1976) Subcultures, Cultures and Class. In Hall, S. and Jefferson, T. (eds.) Resistance Through Rituals. London: Hutchinson.

CLOUT, H.D. (1972) Rural Geography. An Introductory Survey. Oxford: Pergamon Press Ltd.

COKER, P.L. (1978) Syntactic and semantic features in the acquisition of before and after. J. Child Lang., 5, 261-277.

COLE, N. (1981) Bias in Mental Testing. Am. Psychol., 36, 10, 1067-77.

COX, M.V. (1979) Young children's understanding of 'in front of' and 'behind' in the placement of objects. J. Child Lang., 6, 371-374.

COX, M.V. and WILLETTS, E. (1982) Childhood Egocentrism: The Order of Acquisition of Before-Behind and Left-Right Relationships. Br. J. educ. Psychol., 52, 366-369.

CRONBACH, L.J. (1961) Essentials of Psychological Testing. Second Edition. New York: Harper & Row.

CSO (1980) Social Trends 10. London: HMSO.

CSO (1981) Social Trends 11. London: HMSO.

CULLINGFORD, D. and OPENSHAW, S. (1982) Identifying Areas of Rural Deprivation Using Social Area Analysis. Regional Studies, 16, 409-418.

DAHL, T.A. (1973) Test Review of the BTBC. Measurement and Evaluation in Guidance, 6, 1, 63-65.
Also in Buros, O.K. (ed.) (1978) The Eighth Mental Measurements Yearbook, Vol. 1. New Jersey: The Gryphon Press.

DALE, B. (1980) Subjective and Objective Social Indicators in Studies of Regional Social Well-being. Regional Studies, 14, 503-514.

DATTA, L. (1983) Epilogue. In the Consortium for Longitudinal Studies As the Twig is Bent...
Hillsdale, N.J.: Lawrence ERlbaum Assoc.

DAVIE, R., BUTLER, N. and GOLDSTEIN, H. (1972) From Birth to Seven: The second report of the NCDS (1958 cohort). London: Longman.

DAVIS, J. (1974) Performance of Young Hearing-Impaired Children on a Test of Basic Concepts.
J. Speech and Hearing Research, 17, 342-351.

D.E.S. (1967) Children and their Primary Schools. A report of the Central Advisory Council for Education (England) Vol. I: Report (Plowden).
London: HMSO.

D.E.S. (1978) Primary education in England. London: HMSO.

D.E.S. (1980) Statistical Bulletin 17/80. Pupil/Teacher Ratios for each Local Education Authority in England - January 1980.
London: HMSO.

DEPARTMENT OF EMPLOYMENT (1972) Classification of Occupations and Directory of Occupational Titles. London: HMSO.

DONALDSON, M. (1978) Children's Minds. Glasgow: Fontana.

DONALDSON, M. (1982) Conservation: What is the question?
Br. J. Psychol., 73, 199-207.

DONALDSON, M. and BALFOUR, G. (1968) Less is More: A Study of Language Comprehension in Children.
Br. J. educ. Psychol., 59, 461-471.

DONALDSON, M. and McGARRIGLE, J. (1974) Some clues to the nature of semantic development. J. child Lang., 1, 2, 185-194.

DONALDSON, M. and WALES, R. (1970) On the acquisition of some relational terms. In Hayes, J.R. (ed.) Cognition and the Development of Language. New York: Wiley.

DOUGLAS, J.W.B. (1964) The Home and the School. London: MacGibbon and Kee.

DURKIN, K. (1978) Spatial and Temporal Prepositions in the Language of Young Schoolchildren. Unpublished Ph.D.: University of Cambridge.

DURKIN, K. (1981) Aspects of late language acquisition. First Language, 2, 1, 4, 47-59.

EILERS, R.E., OLLER, D.K. and ELLINGTON, J. (1974) The acquisition of word-meaning for dimensional adjectives: the long and the short of it. J. Child Lang., 1, 2, 195-204.

ELKIND, D. (1973) The Case for Academic Preschool: Fact or Fiction. In Smart, M.S. and Smart, R.C. (eds.) Preschool Children: Development and Relationships. New York: Macmillan.

ELLIOTT, C.D. (1983) The British Ability Scales: Technical Handbook. Windsor: NFER/Nelson.

EMLER, N. and VALIANT, G.L. (1982) Social interaction and cognitive conflict in the development of spatial coordination skills. Br. J. Psych., 73, 295-303.

ENGLISH, J., MADIGAN, R. and NORMAN, P. (1976) Slum Clearance. London: Croom Helm.

ERIKSON, E.H. (1967) Growth and Crises of the Healthy Personality. In Lazarus, R.S. and Opton, E.M. (eds.) Personality. Harmondsworth: Penguin.

ERNHARDT. C.B., SPANER, S.D. and JORDAN, T.E. (1977) Validity of Selected Preschool Screening Tests. Contemp. educ. Psychol., 2, 78-89.

ESSEN, J., FOGELMAN, K. and HEAD, J. (1978) Childhood Housing Experiences and School Attainment.
Child: care, health and development, 4, 1, 41-58.

ESSEN, J. and WEDGE, P. (1982) Continuities in Childhood Disadvantage. Studies in Deprivation and Disadvantage 6. SSRC/DHSS. London: Heinemann Educational Books.

ESTES, G., HARRIS, J., MOERS, F. and WODRICH, D. (1976) Predictive validity of the Boehm Test of Basic Concepts for achievement in first grade. Educ. psychol. Measur., 36, 1031-35.

EVANS, A. (1980) Poverty and the Conurbations. In Cameron, G. (ed.) The Future of the British Conurbations. London: Longman Group.

FEAGANS, L. (1980) How to make sense of temporal/spatial 'before' and 'after'. J. Child Lang., 7, 529-537.

FIELD, D. (1981) Can Preschool Children Really Learn to Conserve? Child Dev., 52, 326-334.

FIELD, F. (ed.) (1977) Education and the Urban Crisis. London: Routledge & Kegan Paul.

FISCHER, K.W. (1980) A Theory of Cognitive Development: The Control and Construction of Hierarchies of Skills.
Psychol. Rev., 87, 6, 477-531.

FLAVELL, J.H. (1974) The Development of Inferences About Others. In Mischel, T. (ed.) Understanding Other Persons. Oxford: Blackwell & Mott.

FOXMAN, D.D., RUDDOCK, G.J., BADGER, M.E. and MARTINI, R.M. (1982) Mathematical Development. Primary Survey Report No. 3. London: HMSO.

FRANCIS, R. (1983) Symbols, Images and Social Organisation in Urban Society. In Pons, V. and Francis, R., (eds.) Urban and Social Research: Sociological Review Monograph 30. London: Routledge & Kegan Paul.

FREEMAN, F.S. (1972) Review of the BTBC. In Buros, O.K. (ed.) (1972) The Seventh Mental Measurements Yearbook. New Jersey: The Gryphon Press.

FRENCH, L.A. and BROWN, A.L. (1977) Comprehension of before and after in logical and arbitrary sequences. J. Ch. Lang., 4, 247-256.

FUCHS, E. (1973) How Teachers Learn to Help Children Fail. In Keddie, N. (ed.) Tinker, Tailor... The Myth of Cultural Deprivation. Harmondsworth: Penguin.

GAGNE, R.M. (1970) The Conditions of Learning. Second Edition. London: Holt, Rinehart & Winston.

GANS, H.J. (1968) Urbanism and Suburbanism as Ways of Life. In Pahl, R.E. (ed.) Readings in Urban Sociology. Oxford: Pergamon Press.

GELMAN, R. (1979) Preschool Thought. Am. Psychol., 34, 10, 900-905.

GELMAN, R. (1982) Accessing one-to-one correspondence: Still another paper about conservation. Br. J. Psych., 73, 209-220.

GESELL, A. (1940) The First Years of Life. New York: Harper.

GIBSON, M.S. and LANGSTAFF, M.J. (1982) An Introduction to Urban Renewal. London: Hutchinson.

GITTUS, E. (1978) Deprived Areas in Social Planning. In Herbert, D.T. and Johnston, R.J. (eds.) Social Areas in Cities Vol. II. Chichester: John Wiley & Sons.

GLASS, R. (1968) Urban Sociology in Great Britain. In Pahl, R.E. (ed.) Readings in Urban Sociology. Oxford: Pergamon Press.

GOLDSTEIN, H. (1979) Consequences of using the Rasch model for educational assessment. Br. educ. Res. J., 5, 211-220.

GOLDSTEIN, H. and BLINKHORN, S. (1982) The Rasch model still does not fit. Br. educ. Res. J., 8, 167-170.

GOODACRE, E.J. (1968) Teachers and their Pupils' Home Background. Slough: NFER.

GOODZ, N.S. (1982) Is Before Really Easier to Understand than After? Child Dev., 53, 822-825.

GOPNIK, A. (1982) Words and plans: early language and the development of intelligent action. J. Child Lang., 9, 303-318.

GORMAN, T.P., WHITE, J., ORCHARD, I. and TATE, A. (1982) Language Performance in Schools. Primary Survey Report No. 2. London: HMSO.

GREENO, J.G. (1980) Psychology of Learning, 1960-1980: One Participant's Observations. Am. Psychol., 35, 8. 713-728.

GREENO, J.G. (1982) Response to "The Hegemony of Natural Scientific Conceptions of Learning". Am. Psychol., 37, 332-334.

GRONER, R., KELLER, B. and MENZ, C. (1980) Formal Precision, Where and What for, Or: The Ape Climbs the Tree. In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking. New York: Academic Press.

GUILFORD, J.P. (1982) Cognitive Psychology's Ambiguities: Some Suggested Remedies. Psychol. Rev., 89, 1, 48-59.

HABER, A. and RUNYON, R.P. (1979) General Statistics. Third Edition. London: Addison-Wesley Pub. Co.

HALSEY, A.H. (ed.) (1972) Educational Priority. Vol. I. EPA Problems and Policies. London: HMSO.

HANSON, J. (1983) Out of Touch with Front Doors. The Guardian 27.7.83.

HARGREAVES, D.J., MOLLOY, C.G. and PRATT, A.R. (1982) Social Factors in Conservation. Br. J. Psychol., 73, 231-234.

HARNISCH, D.L. (1983) Item Response Patterns: Applications for Educational Practice. J. educ. Measurement., 20, 2, 191-206.

HARRIS, P. (1974) Inferences and semantic development. J. child Lang., 1, 143-152.

HARRIS, P. (1982) Cognitive prerequisites to language? Br. J. Psychol., 73, 187-195.

HARRIS, P.L., MACRAE, A.J. and BASSETT, E.M. (1978) Disambiguation by Young Children. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

HARTLEY, D. (1980) Sex differences in the Infant School: definitions and 'theories'. Br. J. Sociol. Educ., 1, 1, 93-105.

HERBERT, D.T. and JOHNSTON, R.J. (eds.) (1978) Social Areas in Cities Vol. II. Spatial Perspectives on Problems and Policies. Chichester: John Wiley & Sons.

HINES, B. (1969) A Kestrel for a Knave. Harmondsworth: Penguin.

HODGES, W.L. (1978) The Worth of the Follow Through Experience. Harv. Educ. Rev., 18, 2, 186-191.

HOFF-GINSBURG, E. and SHATZ, M. (1982) Linguistic Input and the Child's Acquisition of Language. Psychol. Bull., 92, 1, 3-26.

HOOGENRAAD, R., GRIEVE, R., BALDWIN, P. and CAMPBELL, R. (1978) Comprehension as an Interactive Process. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

HOOPER, D. and INEICHEN, B. (1979) Adjustment to Moving: A Follow-Up Study of the Mental Health of Young Families in New Housing. Soc. Sci. & Med., 13, 163-168.

HOOPER, F.H., SWINTON, S.S. and SIPPLE, T.S. (1979) Logical Reasoning in Middle Childhood. In Klausmeier, H.J. and Associates. Cognitive Learning and Development: Information-Processing and Piagetian Perspectives. Cambridge, Mass.: Ballinger Pub. Co.

HOUCH, C.K., BISKIN, D.S. and REGETZ, J. (1973) A Comparison of Urban and Rural Reliability Estimates for the Boehm Basic Concept Test. Psychology in the Schools, 10, 4, 430-432.

HOUSE, E.E., GLASS, G.U., McLEAN, L.D. and WALKER, D.F. (1978) No Simple Answer: Critique of the Follow Through Evaluation. Harv. Educ. Rev., 48, 2, 128-160.

HUBBARD, D.N. (1974) Investments in the Preschool Years. Froebel Journal, 30, 37-41.

HUDSON, L.M., GUTHRIE, K.H. and SANTILLI, N.S. (1982) The use of linguistic and non-linguistic strategies in Kindergarteners' interpretation of 'more' and 'less'. J. child Lang., 9, 125-138.

HUGHES, M. and GRIEVE, R. (1980) On Asking Children Bizarre Questions. First Language, 1, 2, 149-160.

HUNT, E. and LANSMAN, M. (1975) Cognitive Theory Applied to Individual Differences. In Estes, W.K. (ed.) Handbook of Learning. Vol. 1. Hillsdale, N.J.: Lawrence Erlbaum Assoc.

HUTT, M.L. (1980) Microdiagnosis and Misuse of Scores and Standards. Psychol. Rep., 50, 239-255.

INEICHEN, B. and HOOPER, D. (1974) Wives' Mental Health and Children's Behaviour Problems in Contrasting Residential Areas. Soc. Sci. & Med., 8, 369-374.

INHELDER, B., SINCLAIR, H. and BOVET, M. (1974) Learning and the Development of Cognition. London: Routledge & Kegan Paul.

ISAACS, S. (1933) Social Development in Young Children. London: Routledge & Kegan Paul.

JACKSON, B. and MARSDEN, D. (1962) Education and the Working Class. London: Routledge & Kegan Paul.

JACKSON, P.M. (1980) The socio-economic impact of urban renewal. In Cameron, G. (ed.) The Future of British Conurbations. London: Longman Group.

JENCKS, C. (1973) Inequality: a reassessment of the effects of family and schooling in America. New York: Harper & Row, Inc.

JENSEN, A.R. (1980) Bias in Mental Testing. London: Methuen & Co.

JERSILD, A.T. (1960) Child Psychology. Fifth Edition. London: Staples Press.

KARMILOFF-SMITH, A. (1978) The Interplay Between Syntax, Semantics and Phonology in Language Acquisition Processes. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

KARNES, M.B. (1973) Evaluation and implications of research with young handicapped and low-income children. In Stanley, J.C. (ed.) Compensatory Education for Children. Ages 2 to 8. London: The John Hopkins University Press.

KARNES, M.B., SHWEDEL, A.M. and WILLIAMS, M.B. (1983) A Comparison of Five Approaches for Educating Children from Low-Income Homes. In The Consortium for Longitudinal STudies. As the Twig is Bent... Hillsdale, N.J.: Lawrence Erlbaum Assoc.

KAUFMAN, A.S. (1978) The Importance of Basic Concepts in the Individual Assessment of Preschool Children. J. School Psychol., 16, 3, 207-211.

KAVANAUGH, R.D. (1979) Observations on the role of logically constrained sentences in the comprehension of 'before' and 'after'. J. Child Lang., 6, 353;357.

KEATS, D.M. and KEATS, J.A. (1978) Evidence Significant in the Language-Though Relationship. In Keats, J.A., Collis, K.F. and Halford, G.S. (eds.) Cognitive Development. Research Based on a Neo-Piagetian Approach. Chichester: John Wiley & Sons.

KELSALL, R.K. and KELSALL, H.M. (1971) Social Disadvantage and Educational Opportunity. London: Holt, Rinehart & Winston.

KENDLER, H.H. and GUENTHER, K. (1980) Developmental Changes in Classificatory Behaviour. Child Dev., 51, 339-348.

KENDLER, H.H. and KENDLER, T.S. (1975) From discrimination learning to cognitive development: a new-behavioristic odyssey. In Estes, W.K. (ed.) Handbook of Learning and Cognitive Processes. Vol. I. Hillsdale, N.J.: Lawrence Erlbaum Assoc.

KIRK, S.A., McCARTHY, J.J. and KIRK, W.D. (1968) Examiners Manual. Illinois Test of Psycholinguistic Abilities. Urbana: University of Illinois Press.

KLAHR, D. (1980) Information-Processing Models of Intellectual Development. In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking. New York: Academic Press.

KLAUSMEIER, H.J. (1979a) Introduction In Klausmeier, H.J. and Associates Cognitive Learning and Development: Information-Processing and Piagetian Perspectives. Cambridge, Mass: Ballinger Pub. Co.

KLAUSMEIER, H.J. (1979b) Organization of Cognitive Abilities Based on CLD Theory. In Klausmeier, H.J. and Associates Cognitive Learning and Development: Information-Processing and Piagetian Perspectives. Cambridge, Mass.: Ballinger Pub. Co.

KNOX, P.L. (1972) Spatial Variations in the Level of Living in England and Wales. Unpublished Ph.D.: University of Sheffield.

KNOX, P.L. (1975) Social Well-Being: A Spatial Perspective. London: Oxford University Press.

KUCZAJ, S. and MARATSOS, M. (1975) On the acquisition of front, back and side. Child Dev., 46, 202-210.

LABOV, W. (1969) The Logic of Nonstandard English. In Giglioli, P.P. (ed.) Language and Social Context. Harmondsworth: Penguin.

LAWLOR, G. (1972) Review of the BTBC. In Buros, O.K. (ed.) The Seventh Mental Measurements Yearbook. New Jersey: The Gryphon Press.

LAWTON, J.T. and HOOPER, F.H. (1978) Piagetian Theory and Early Childhood Education: A Critical Analysis. In Siegel, L.S. and Brainerd, C.J. (eds.) Alternatives to Piaget. New York: Academic Press.

LAXON, V.J. (1981) On the Problems of Being More or Less the Same. J. exp. Child Psychol., 31, 531-543.

LAZAR, I. (1983) Discussion and Implications of the Findings. In The Consortium for Longitudinal Studies. As the Twig is Bent... Hillsdale, N.J.: Lawrence Erlbaum Assoc.

LEINHARDT, G. and SEEWALD, A.M. (1981) Overlap: What's Tested, What's Target? J. educ. Measurement, 18, 2, 85-95.

LEVIN, G.H., HENDERSON, B., LEVIN, A.M. and HOFFER, G.L. (1975) Measuring Knowledge of Basic Concepts by Disadvantaged Preschoolers. Psychol. in the Schs., 12, 132-139.

LEVINE, S.C. and CAREY, S. (1982) Up front: the acquisition of a concept and a word. J. Child Lang., 9, 645-657.

LIGHT, P.H. (1983) Social interaction and cognitive development: a review of post-Piagetian research. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

LIGHT, P.H., BUCKINGHAM, N. and ROBBINS, A.H. (1979) The Conservation Task as an Interactional Setting. Br. J. educ. Psychol., 49, 3, 304-310.

LINN, R.L. and HARNISCH, D.L. (1981) Interactions Between Item Content and Group Membership on Achievement Test Items. J. educ. Measurement, 18, 2, 109-118.

LITTLE, A. (1977) What is happening in the inner city schools? In Field, F. (ed.) Education and the Urban Crisis. London: Routledge & Kegan Paul.

LUNT, I. (1983) Assessment of Cognitive Development. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

LUNZER, E.A., WILKINSON, J.E. and DOLAN, T. (1976) The Distinctiveness of Operativity as a Measure of Cognitive Functioning in Five-Year-Old Children. Br. J. educ. Psychol., 46, 3, 280-94.

LURIA, A.R. (1961) The Role of Speech in the Regulation of Normal and Abnormal Behaviour. Oxford: Pergamon Press.

LUSZCZ, M.A. and BACHARACH, V.R. (1980) Preschoolers' Picture Recognition Memory: The Pitfalls of Knowing How a Thing Shall Be Called. Canad. J. Psychol., 34, 2, 155-160.

MACAULEY, R.S. (1978) The myth of female superiority in language. J. Child Lang., 5, 353-363.

McCANDLESS, B.R. (1972) Review of the BTBC. In Buros, O.K. (ed.) The Seventh Mental Measurements Yearbook. New Jersey: The Gryphon Press.

McKENZIE, M. (1980) Education and the British Conurbations. In Cameron, G. (ed.) The Future of the British Conurbations. London: Longman Group.

MACNAMARA, J. (1977a) Problems about Concepts. In Macnamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

MACNAMARA, J. (1977b) From Sign to Language. In Macnamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

McNEIL, D. (1982) Explicating the Conceptual Basis of Language. J. Child Lang., 9, 521-526.

McNEMAR, Q. (1962) Psychological Statistics. Third Edition. New York: John Wiley & Sons, Inc.

MANNING, K. and SHARP, A. (1977) Structuring Play in the Early Years at School. Schools Council Project. London: Ward Lock Educational/Drake Educational Associates.

MARSLEN-WILSON, W.D. and TYLER, L.K. (1981) Central processes in speech understanding. In Longuet-Higgins, H.C., Lyons, J. and Broadbent, D.E. (eds.) The Psychological Mechanisms of Language. London: The Royal Society and The British Academy.

MARTLEW, M. (1980) The Effect of Adult Role on the Language used by Preschool Children. In First Language, 1, 3, 3, 191-207.

MATTHEWS, G. and MATTHEWS, J. (1978) Early Mathematical Experience. London: Schools Council/Addison-Wesley Ltd.

MEADOWS, S. (1983) As assessment of Piaget's theory of cognitive development. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

MEISSNER, J.A. (1975) Use of relational concepts by inner city children. J. educ. Psychol., 67, 22-29.

MIDWINTER, E. (1969) Educational Priority Areas: The Philosophic Question. Occasional Papers No. 1. Liverpool: Liverpool EPA Project.

MIDWINTER, E. (1970) Home and School Relations in Educational Priority Areas. Occasional Papers No. 4. Liverpool: Liverpool EPA Project.

MILLER, L.B. and BIZZELL, R.P. (1983) Long-Term Effects of Four Preschool Programs: Sixth, Seventh and Eighth Grades. In Child Dev., 54, 727-741.

MILLER, S.A. (1982) On the generalizability of conservation: A comparison of different kinds of transformation. In Br. J. Psychol., 73, 221-230.

MILLS, M. and FUNNELL, E. (1983) Experience and cognitive processing. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

MOERS, F. and HARRIS, J. (1978) Instruction in Basic Concepts and First-Grade Achievement. Psychol. in the Schs., 15, 1, 84-86.

MOORE, T.E. and HARRIS, A.E. (1978) Language and Thought in Piagetian Theory. In Siegel, L.S. and Brainerd, C.J. (eds.) Alternatives to Piaget: critical essays on the theory. New York: Academic Press.

MURIE, A. (1983) Housing Inequality and Deprivation. SSRC/DHSS Studies in Deprivation and Disadvantage. London: Heinemann Educational.

NASH, R. (1973) Classrooms Observed. The Teacher's Perception and the Pupil's Performance. London: Routledge & Kegan Paul.

NELSON, K. (1977) The Conceptual Basis for Naming. In Macmanara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

NELSON, K. (1978) Semantic Development and the Development of Semantic Memory. In Nelson, K.E. (ed.) Children's Language. Vol. I. New York: Gardner Press Inc.

NELSON, K.E. and BONVILLIAN, J.D. (1978) Early Language Development: Conceptual Growth and Related Processes Between 2 and 4½ Years of Age. In Nelson, K.E. (ed.) Children's Language. Vol. I. New York: Gardner Press Inc.

NELSON, K.E. and NELSON, K. (1978) Cognitive Pendulums and the Linguistic Realization. In Nelson, K.E. (ed.) Children's Language. Vol. I. New York: Gardner Press Inc.

NELSON, R.B. and CUMMINGS, J.A. (1981) Basic Concept Attainment of Educable Mentally Handicapped Children: Implications for Teaching Concepts. Education and Training of the Mentally Retarded. Dec., 303-306.

NEWSON, J. and NEWSON, E. (1965) Patterns of Infant Care in an Urban Community. Harmondsworth: Penguin (Pelican) (originally pub. 1963 Allen & Unwin.)

NEWSON, J. and NEWSON, E. (1968) Four Years Old in an Urban Community. London: George Allen & Unwin.

NEWSON, J. and NEWSON, E. (1978) Seven Years Old in the Home Environment. Harmondsworth: Penguin (Pelican) (originally pub. 1976 Allen & Unwin.)

NOLL, V.H. (1972) Review of the BTBC. In Buros, O.K. (ed.) The Seventh Mental Measurements Yearbook. New Jersey: The Gryphon Press.

NUFFIELD FOUNDATION (1967, 1970) Nuffield Mathematics Project. London: W. & R. Chambers and John Murray.

O'BRYAN, K.G. (1975) Orientations in Compensatory Education. In Wedell, K. (ed.) Orientations in Special Education. London: Johh Wiley & Sons.

OLSON, D.R. (1977) The Contexts of Language Acquisition. In Macnamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

OLSON, D.R. and TORRANCE, N.G. (1983) Literacy and cognitive development: a conceptual transformation in the early school years. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

OPCS (1973) Census 1971. England and Wales. County Report. Yorkshire West Riding. Part I. London: HMSO.

OPCS (1980) Classification of Occupations. London: HMSO.

OSBORNE, J.W. (1982) The Hegemony of Natural Scientific Conceptions of Learning. In Am. Psychol., 37, 330-332.

PAHL, R.E. (1968) The Rural-Urban Continuum. In Pahl, R.E. (ed.) Readings in Urban Sociology. Oxford: Pergamon Press.

PARRY, M. and ARCHER, H. (1974) Pre-School Education. London: Schools Council/Macmillan.

PARRY, M. and ARCHER, H. (1975) Two to five. London: Schools Council/Macmillan.

PASCUAL-LEONE, J. (1980) Constructive Problems for Constructive Theories: The Current Relevance of Piaget's Work and a Critique of Information-Processing Simulation Psychology. In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking. New York: Academic Press.

PELLOW, D. (1981) The New Urban Community: Mutual Relevance of the Social and Physical Environments. Human Organization, 40, 1, 15-26.

PIAGET, J. (1952) The Child's Conception of Number. London: Routledge & Kegan Paul.

PIAGET, J. (1955) Language and Thought of the Child. London: Routledge & Kegan Paul.

PIAGET, J. (1977a) The Development of Thought. London: Routledge & Kegan Paul.

PIAGET, J. (1977b) The Grasp of Consciousness. London: Routledge & Kegan Paul.

PIAGET, J. (1978) Success and Understanding. London: Routledge & Kegan Paul.

PIAGET, J. and INHELDER, B. (1956) The Child's Conception of Space. London: Routledge & Kegan Paul.

PIDGEON, D.A. (1970) Expectation and Pupil Performance. Slough: NFER.

PIERSEL, W.C. and McANDREWS, T. (1982) Concept Acquisition and School Progress: An Estimation of the Boehm Test of Basic Concepts. Psychol. Rep., 50, 738-786.

PIERSEL, W.C. and REYNOLDS, C.R. (1981) Factorial Validity of Item Classification on the Boehm Test of Basic Concepts (BTBC), Forms A and B. Educ. psychol. Measur., 41, 579-583.

PILLING, D. and PRINGLE, M.K. (1978) Controversial Issues in Child Development. London: Paul Elek Ltd.

PLOMIN, R. and FOCH, T.T. (1981) Sex Differences and Individual Differences. Child Dev., 52, 383-85.

POULTON, G.A. and JAMES, T. (1975) Pre-School Learning in the Community. London: Routledge & Kegan Paul.

PRAWAT, R.S. and WILDFONG, G.S. (1980) The Influence of Functional Context on Children's Labeling Responses. Child Dev., 51, 1057-60.

PRINGLE, M.L.K. (1974) The Needs of Children. London: Hutchinson.

PROGER, B.B. (1972) Review of the BTBC. In Buros, O.K. (ed.) The Seventh Mental Measurements Yearbook. New Jersey: The Gryphon Press.

PYLYSHYN, Z.W. (1977a) Children's Internal Descriptions. In Macnamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

PYLYSHYN, Z.W. (1977b) What Does it Take to Bootstrap a Language? In Macnamara, J. (ed.) Language, Learning and Thought. New York: Academic Press.

REYNOLDS, C.R. and PIERSEL, W.C. (1983) Multiple aspects of bias on the Boehm Test of Basic Concepts (Forms A and B) for white and for Mexican-American Children. J. Psychoeducational Assessment, 1, 135-142.

RICHARDS, M.M. (1979a) Sorting out what's in a word from what's not: evaluating Clark's semantic features acquisition theory. J. exp. Child Psychol., 27, 1-47.

RICHARDS, M.M. (1979b) Adjective ordering in the language of young children: an experimental investigation. J. Child Lang., 6, 2, 253-297.

ROBINSON, E. (1983) Metacognitive Development. In Meadows, S. (ed.) Developing Thinking. London: Methuen.

RODGON, M.M. (1978) Overt Action and the Semantic Expression of Action in One-Word Utterances. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

ROMMETVEIT, R. (1978) One the Relationship Between Children's Mastery of Piagetian Cognitive Operations and their Semantic Competence. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

ROSEN, H. (1972) Language and Class. London: Falling Wall Press.

ROSENBACH, J.H. and MOWDER, B.A. (1981) Test Bias: The Other Side of the Coin. Psychol. in the Schs., 18, 450-454.

ROSENSHINE, B. and FURST, N. (1973) The Use of Direct Observation to Study Teaching. In Travers, R.M.W. (ed.) Second Handbook of Research on Teaching. Chicago: Rand McNally & Co.

ROSENTHAL, R. and JACOBSON, L. (1968) Pygmalion in the Classroom. Teacher Expectation and Pupils' Intellectual Development. New York: Holt, Rinehart & Winston.

RUSSELL, J. (1982) The child's appreciation of the necessary falsehood of propositions. Br. J. Psychol., 73, 253-266.

RUTTER, M.D. (1981) The City and the Child. Am. J. Orthopsychiat., 51, 4, 611-625.

RUTTER, M.D. and QUINTON, D. (1977) Psychiatric Disorder - Ecological factors and concepts of causation. In McGurk, H. (ed.) Ecological Factors in Human Development. Amsterdam: North-Holland Pub. Co.

SCHEUNEMAN, J. (1979) A Method of Assessing Bias in Test Items. J. educ. Measurement, 16, 3, 143-152.

SCHOOLS COUNCIL (1964) An Introduction to Some Techniques of Examining. Examinations Bulletin No. 3. London: HMSO.

SCHOOLS COUNCIL (1972, 1974) Science 5/13.
London: Macdonald Educational.

SHAFER, R.E. (1978) A Cross-National Study of Teacher Language Attitudes in England and the United States. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

SHIELDS, M.M. and STEINER, E. (1973) The Language of 3-5 Year Olds in Preschool Education. Educ. Res., 15, 2, 97-105.

SIEGEL, L.S. (1978) The Relationship of Language and Thought in the Preoperational Child. In Siegel, L.S. and Brainerd, C.J. (eds.) Alternatives to Piaget: critical essays on the theory. New York: Academic Press.

SIGSWORTH, A. (1980) The Rural School - in, of, or for the Community? Three Perspectives of Educational Disadvantage. In Centre for Information and Advice on Educational Disadvantage. Educational Disadvantage in Rural Areas.
Manchester: Centre for Educational Disadvantage.

SILVERSTEIN, A.B., BELGER, K.A. and MORITA, D.N. (1982) Social Class Differences on the Boehm Test of Basic Concepts: Are They Due to Bias? Psychol. in the Schs., 19, 431-432.

SINHA, C. and WALKERDINE, V. (1974) Spatial and Temporal Relations in the Linguistic and Cognitive Development of Young Children.
Mimeo Report: School of Education, University of Bristol.

SIPPLE, T.S. and KLAUSMEIR, H.J. (1979) Design and Methods.
In Klausmeier, H.J. and Associates. Cognitive Development: Information-Processing and Piagetian Perspectives.
Cambridge, Mass.: Balinger Pub. Co.

SLOBIN, D.I. (1972) Seven Questions about Language Development.
In Dodwell, P.C. (ed.) New Horizons in Psychology 2.
Harmondsworth: Penguin.

SMEDSLUND, J. (1977) Piaget's Psychology in Practice.
Br. J. educ. Psychol., 47, 1-6.

SMITH, E.F. (1977) An Analysis of Teacher Effectiveness in Nursery Classes.
Unpublished M.Ed. thesis: University of Manchester.

SMITH, E.F. (1980) Organisation in the Nursery.
Education 3-13, 8, 2, 41-43.

SMITH, G. (1975) Educational Priority. Vol. 4: The West Riding EPA.
London: HMSO.

SMOCK, C.D. (1972) Review of the BTBC. In Buros, O.K. (ed.) The Seventh Mental Measurements Yearbook. New Jersey: Gryphon Press.

SOUTH YORKSHIRE C.C. (1980) South Yorkshire Statistics.
Barnsley: S.Y.C.C., County Hall.

SPADA, H. and KLUWE, R.H. (1980) Two Models of Intellectual Development and Their Reference to the Theory of Piaget. In Kluwe, R.H. and Spada, H. (eds.) Developmental Models of Thinking.
New York: Academic Press.

SPECTOR, C.C. (1979) The Boehm Test of Basic Concepts: Exploring the Test Results for Cognitive Deficits.
J. Learn. Disabil., 12, 8, 66-69.

STEINBAUER, E. and HELLER, M.S. (1978) The Boehm Test of Basic Concepts as a predictor of academic achievement in grades 2 and 3.
Psychol. in the Schs., 15, 357-360.

STEINERT, M.C. (1978) Construct and Criterion - Related Validity of the Boehm Test of Basic Concepts.
Unpublished Ph.D.: Kent State University.

STUBBS, M. (1976) Language, Schools and Classroom. London: Methuen.

TAIT, M. and ROBERTS, M. (1974) Play, Language and Experience.
London: OMEP, 24, Woburn Square.

THE CONSORTIUM FOR LONGITUDINAL STUDIES (1983) As the Twig is Bent... Lasting Effects of Preschool Programs. Hillsdale, N.J.: Lawrence Erlbaum Assoc.

TIZARD, B. (1974) Pre-School Education in Great Britain. A Research Review. London: SSRC.

TOMIKAWA, S.A. and DODD, D.H. (1980) Early Word Meanings: Perceptually or Functionally Based? Child Dev., 51, 1103-9.

TOMLINSON-KEASEY, C., EISERT, D.C., KAHLE, L.R. HARDY-BROWN, K. and KEASEY, B. (1979) The Structure of Operational Thought. In Child Dev., 50, 1153-63.

TORREY, J. (1973) Illiteracy in the Ghetto. In Keddie, N. (ed.) Tinker, Tailor... The Myth of Cultural Deprivation. Harmondsworth: Penguin.

TOUGH, J. (1970) Some Differences in the Use of Language in Groups of Three Year Old Children. Unpublished paper: University of Leeds Institute of Education.

TOUGH, J. (1973a) Focus on Meaning. London: Unwin Education Books.

TOUGH, J. (1973b) The Language of Young Children. Lecture given at the University College of Swansea.

TOUGHT, J. (1976) Listening to Children Talking: A guide to the appraisal of children's use of language. Schools Council Communication Skills in Early Childhood Project. London: Ward Lock/Drake Education Assoc.

TOWNSEND, D.J. (1976) Do children interpret "marked" comparative adjectives as their opposites? J. Child Lang., 3, 385-396.

TOWNSEND, D.J. and ERB, M. (1975) Children's strategies for interpreting complex questions. J. Child Lang., 2, 2, 271-277.

TREHUB, S.E. and ABRAMOVITCH, R. (1978) Less is not more: Further observations on non-linguistic strategies. J. exp. child Psychol., 25, 160-167.

TROSBORG, A. (1982) Children's comprehension of 'before' and 'after' reinvestigated. J. Child Lang., 9, 381-402.

TRUDGILL, P. (1974) Sociolinguistics. Harmondsworth: Penguin.

VERNON, M.D. (1968) Perceptions and Perceptual Learning. In Lunzer, E.A. and Morris, J.F. (eds.) Development in Human Learning. London: Stapes Press.

VUYK, R. (1981a, 1981b) Overview and Critique of Piaget's Genetic Epistemology 1965-1980. Vol. 1 and Vol. 2. London, Academic Press.

VYGOTSKY, L.S. (1962) Thought and Language. London: John Wiley & Sons.

WARD, C. (1977) The Child in the City. London: The Architectural Press Ltd.

WEBSTER, N. (1947) Webster's Collegiate Dictionary. Fifth Edition. London: G. Bell & Sons Ltd.

WEDGE, P. and PROSSER, H. (1973) Born to Fail? London: Arrow Books.

WEIL, J. and STENNING, K. (1978) A comparison of young children's comprehension and memory for statements of temporal relations. In Campbell, R.N. and Smith, P.T. (eds.) Recent Advances in the Psychology of Language. Vol. I. London: Plenum Press.

WELLS, G. (1981) Some Antecedents of Early Educational Attainment. Br. J. Sociol., 2, 2, 181-199.

WILCOX, S. and PALERMO, D.S. (1974) "In", "on", and "under" revisited. Cognition. 3, 245-254.

WILCOX, S. and PALERMO, D.S. (1982) Information in Commands. J. Child Lang., 9, 139-150.

WILKENING, F. and ANDERSON, N.H. (1982) Comparison of Two Rule-Assessment Methodologies for Studying Cognitive Development and Knowledge Structure. In. Psychol. Bull., 92, 1, 215-237.

WILKINSON, E., GRANT, D. and WILLIAMSON, D.J. (1978) Strathclyde Experiment in Education. Govan Project. A Public Report. Unpublished paper: University of Glasgow, Dept. of Education.

WILSON, H. and HERBERT, G.W. (1978) Parents and Children in the Inner City. London: Routledge & Kegan Paul.

WINNICOTT, D.W. (1974) Playing and Reality. Harmondsworth: Penguin.

WISEMAN, S. (1967) The Manchester Survey. App. 4, Plowden Report. In Central Advisory Council for Education (England) Children and their Primary Schools. Vol. II. London: HMSO.

WISLER, C.E., BURNS, G.P. and IWAMOTO, D. (1978) Follow Through Redux: A Response to the Critique by House, Glass, McLean and Walker. In Harv. Educ. Rev., 48, 2, 171-185.

WOHWILL, J.F. and HEFT, H. (1977) Environments Fit for the Developing Child. In McGurk, H. (ed.) Ecological Factors in Human Development. Amsterdam: North-Holland Pub. Co.

WOOD, R. (1978) Fitting the Rasch model - A heady tale. Br. J. maths. statist. Psychol., 31, 27-32.

WOODHEAD, M. (1976) Intervening in Disadvantage: A Challenge for Nursery Education. Windsor: NFER Pub. Co. Ltd.

WRIGHT, B.D. (1977) Solving measurement problems with the Rasch model. J. educ. Measurement, 14, 97-116.

YOUNG, A.W. and McPHERSON, J. (1976) Ways of Making Number Judgments and Children's Understanding of Quantity Relations. Br. J. educ. Psychol., 46, 328-332.

YOUNG, M. and WILLMOTT, P. (1957) Family and Kinship in East London. London: Routledge & Kegan Paul.

ZAMMARELLI, J. and BOLTON, N. (1977) The Effects of Play on Mathematical Concept Formation. Br. J. educ. Psychol., 47, 155-161.

ZIGLER, E.F. and BUTTERFIELD, E.L. (1968) Motivational aspects of change in I.Q. test performance of culturally deprived nursery school children. Child Dev., 39, 1-14.

Appendix 1

BTBC Concepts (1971 Edition)

Booklet 1

1. Top
2. Through
3. Away
4. Next to
5. Inside
6. Some, not many
7. Middle
8. Few
9. Farthest
10. Around
11. Over
12. Widest
13. Most
14. Between
15. Whole
16. Nearest
17. Second
18. Corner
19. Several
20. Behind
21. Row
22. Different
23. After
24. Almost
25. Half

Booklet 2

26. Centre
27. As many
28. Side
29. Beginning
30. Other
31. Alike
32. Not first or last
33. Never
34. Below
35. Matches
36. Always
37. Medium-sized
38. Right
39. Forward
40. Zero (without any)[1]
41. Above
42. Every
43. Separated
44. Left
45. Pair
46. Skip (miss)[1]
47. Equal
48. In order
49. Third
50. Least

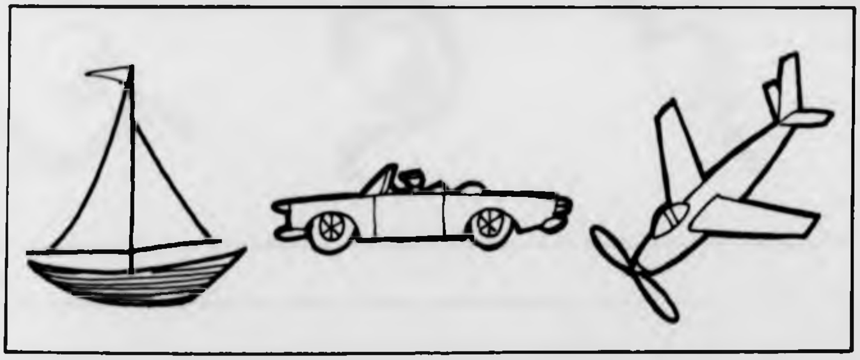
[1] Concepts in parenthesis are those used in the present adapted version.

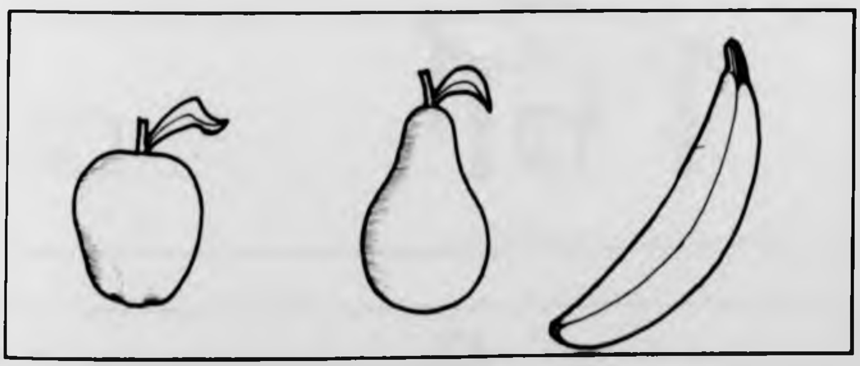
Appendix 2

The BTBC Form A, Booklets 1 and 2.

Form A
Booklet 1

NAME ____________________





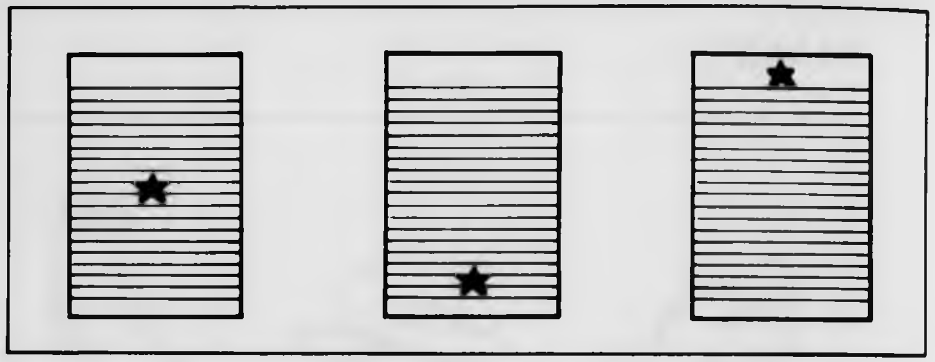
BOEHM TEST of basic concepts

Ann E. Boehm

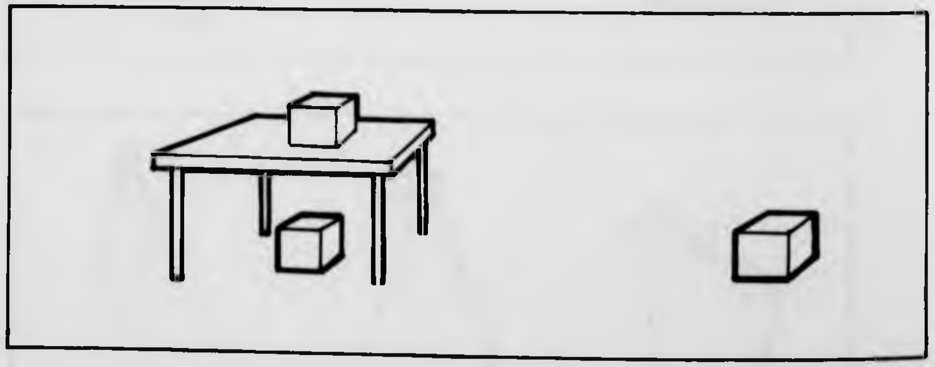
THE PSYCHOLOGICAL CORPORATION, 304 East 45th Street, New York, N.Y. 10017

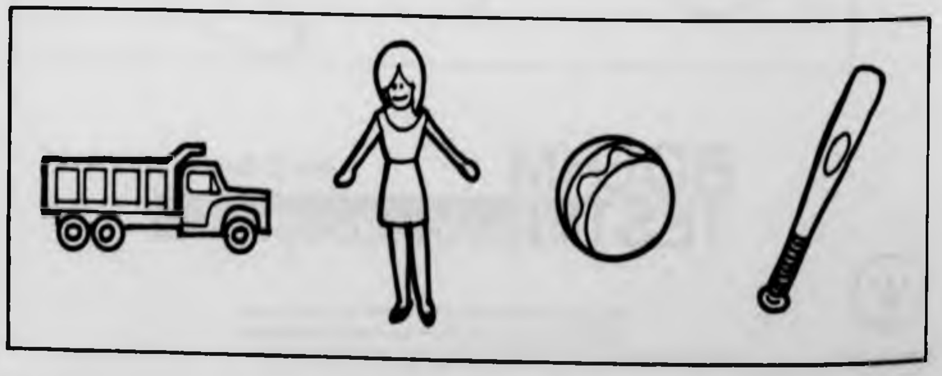
Printed in U.S.A.

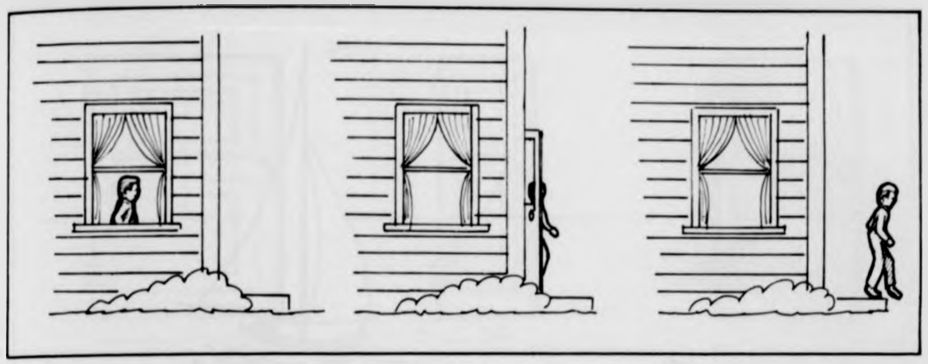
70-182T

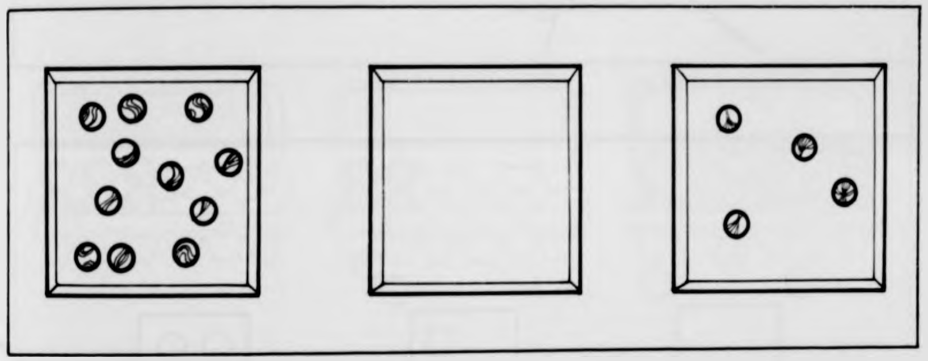

1

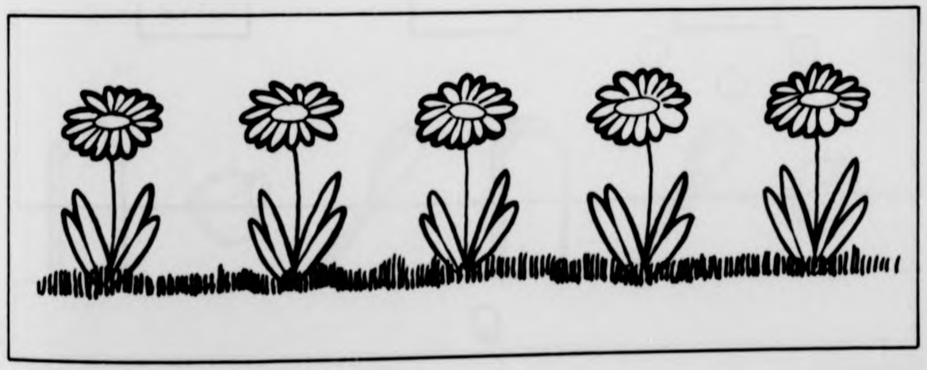

2

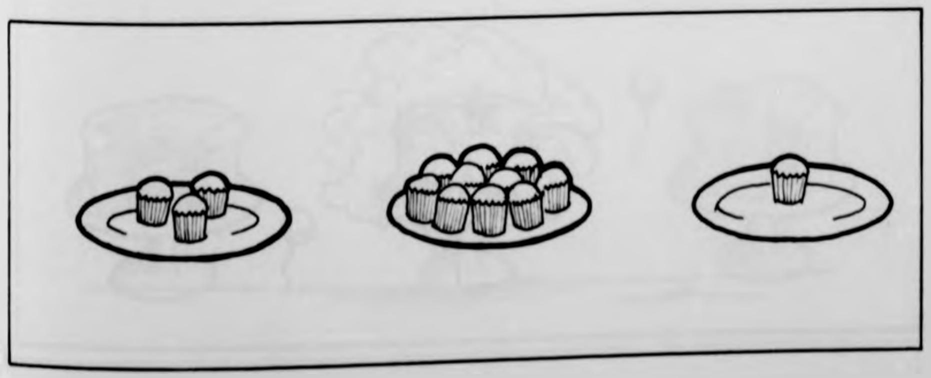

3

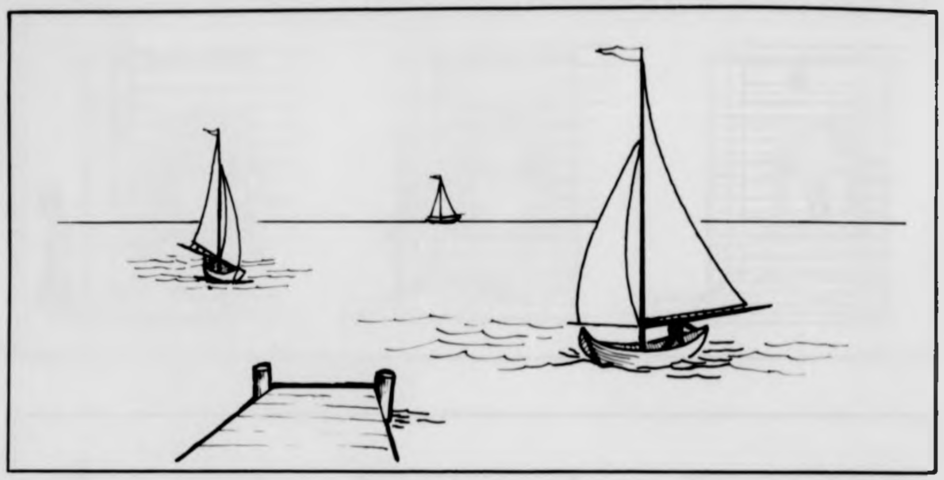

4


5

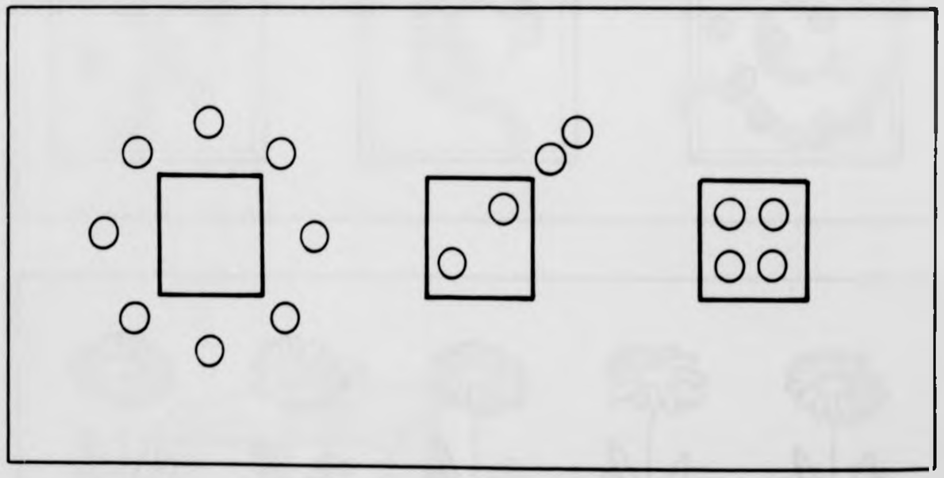

6


7

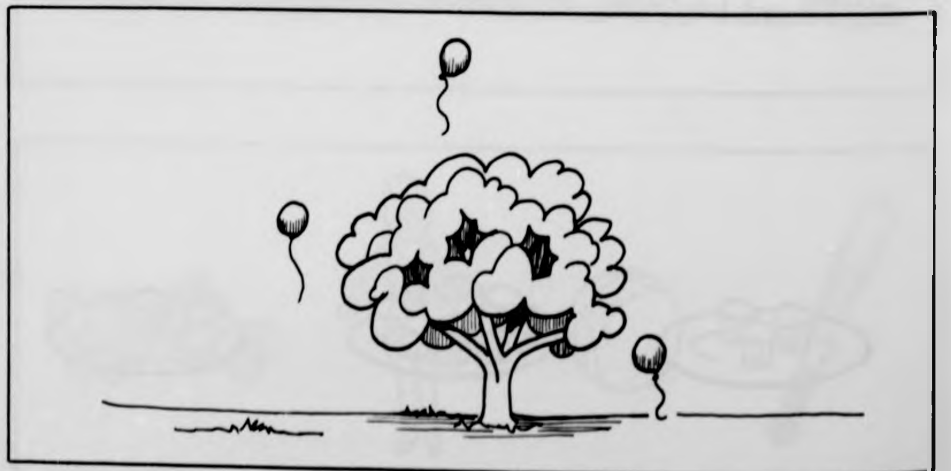

8



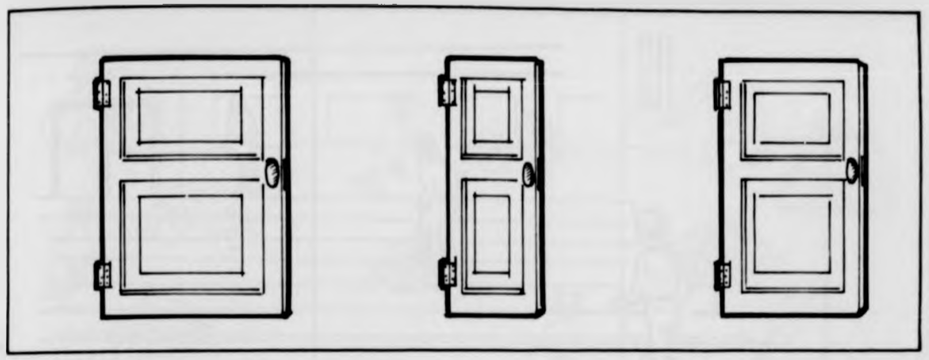
9



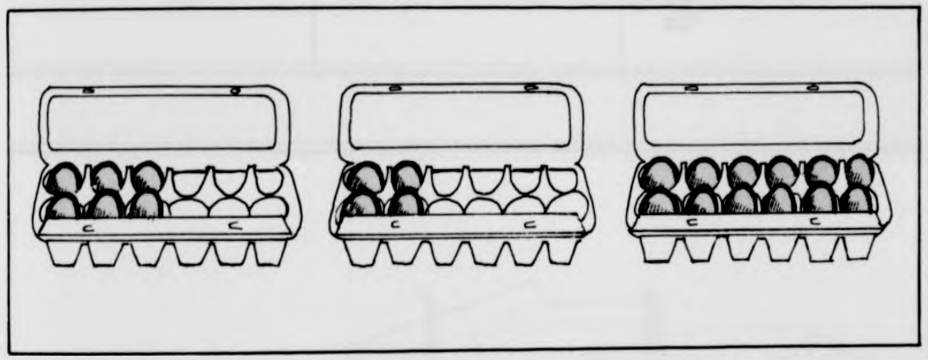
10



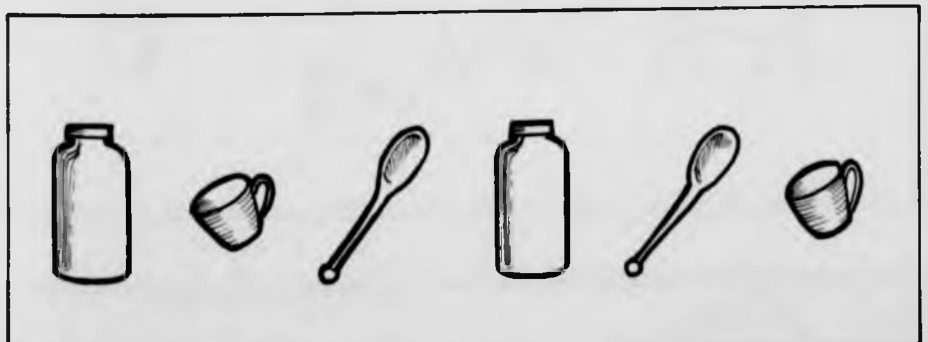
11



12

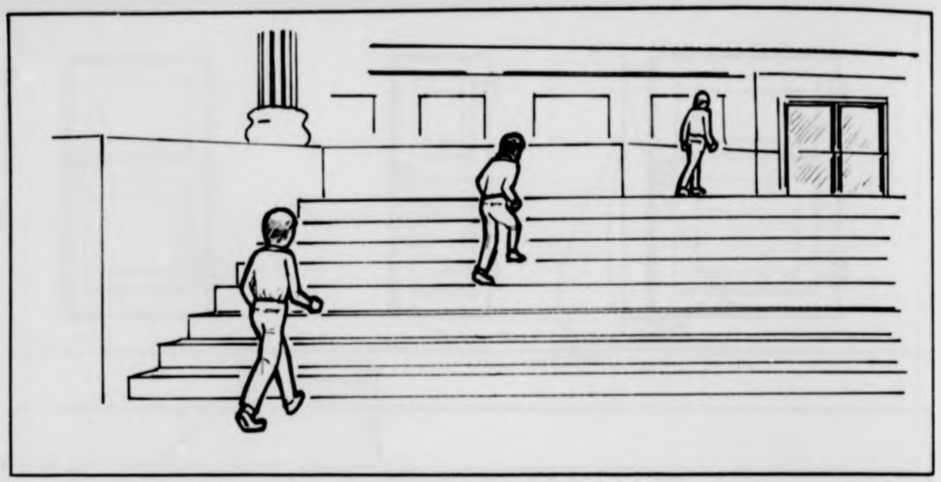


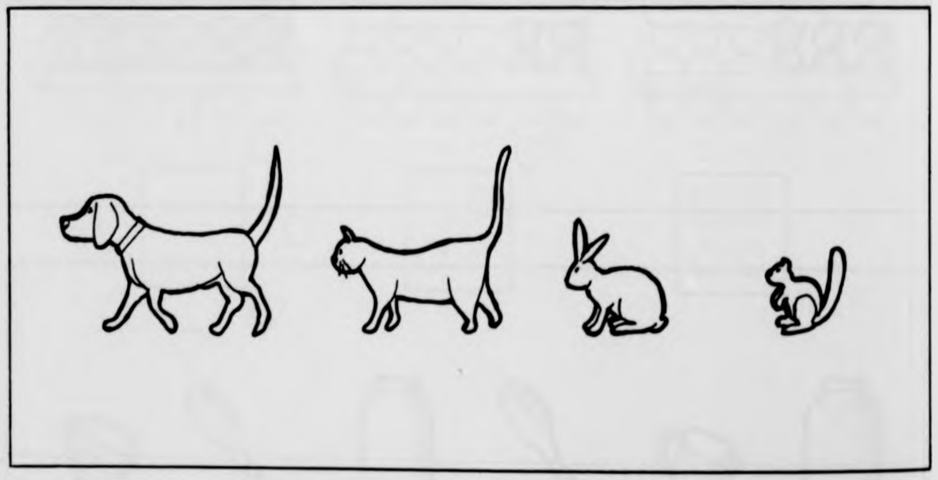
13



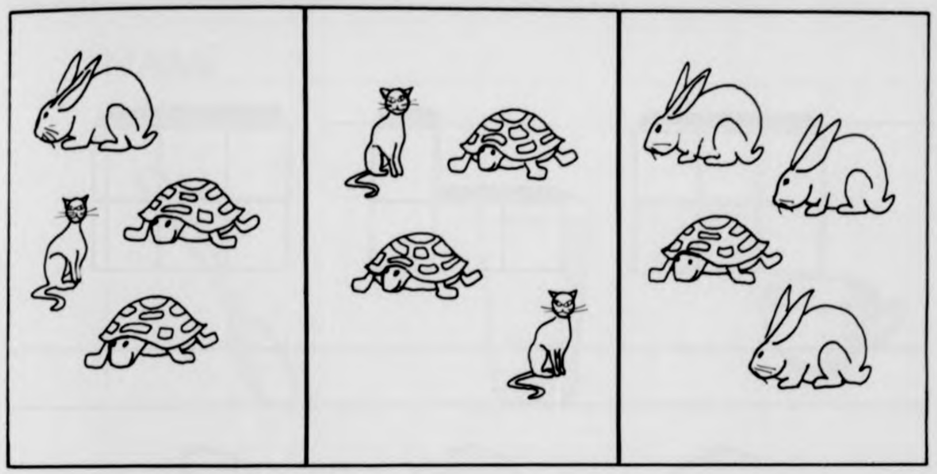
14

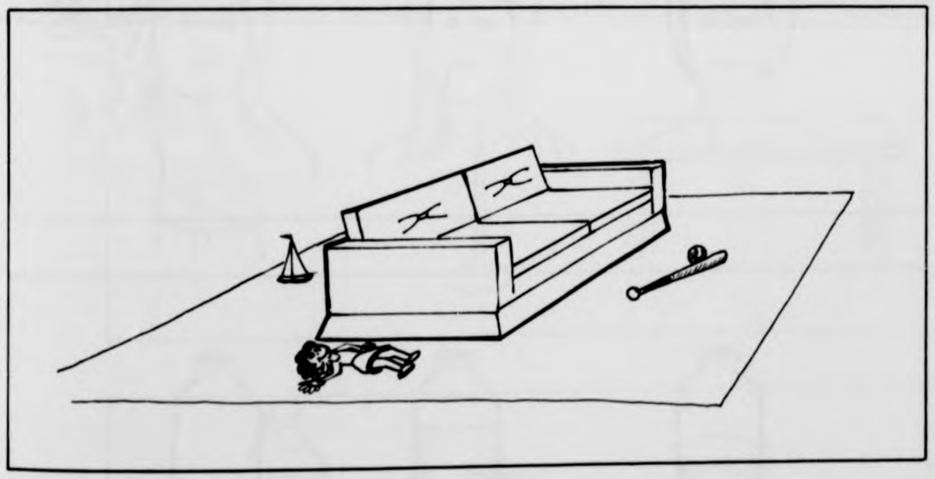
15

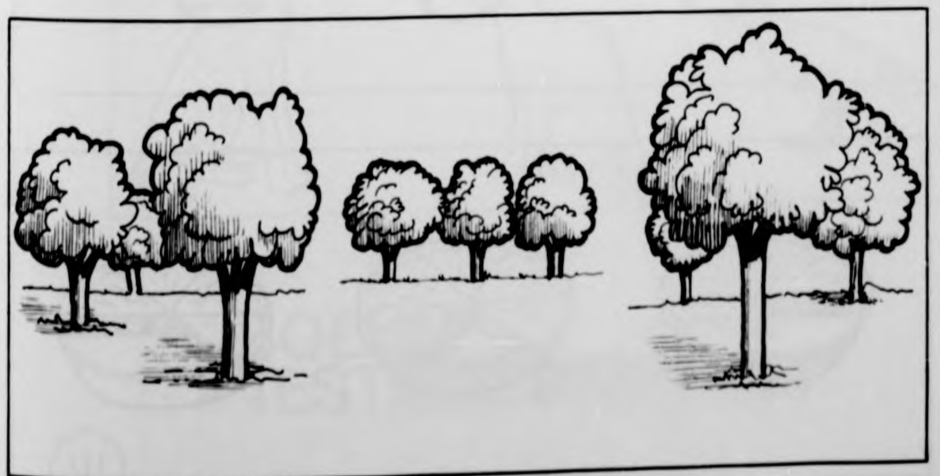

16

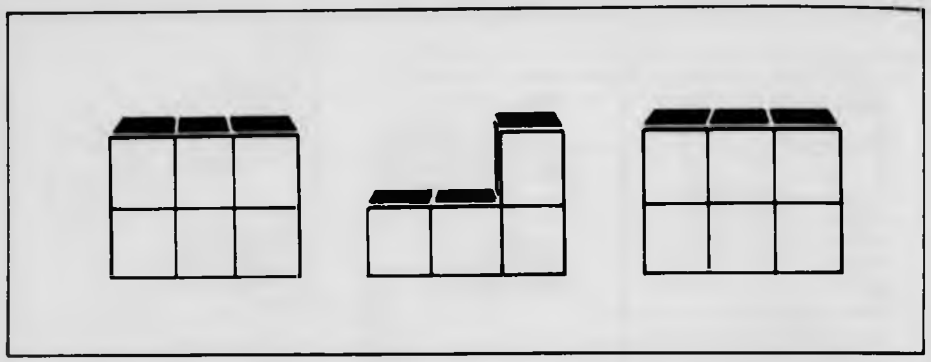

17


18

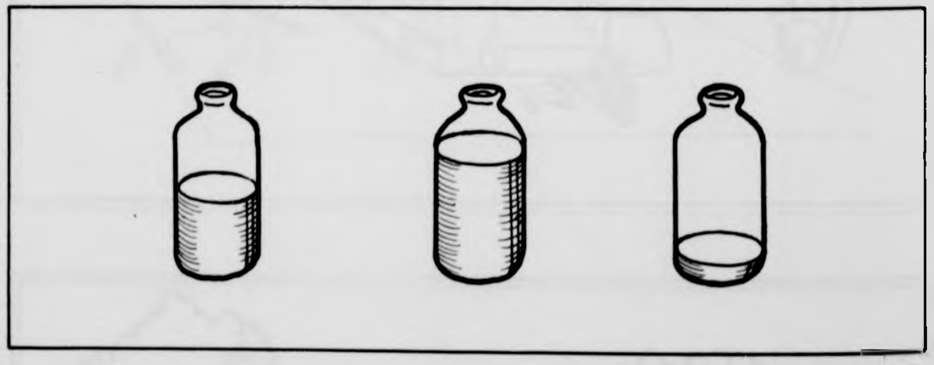

19


20


21


22

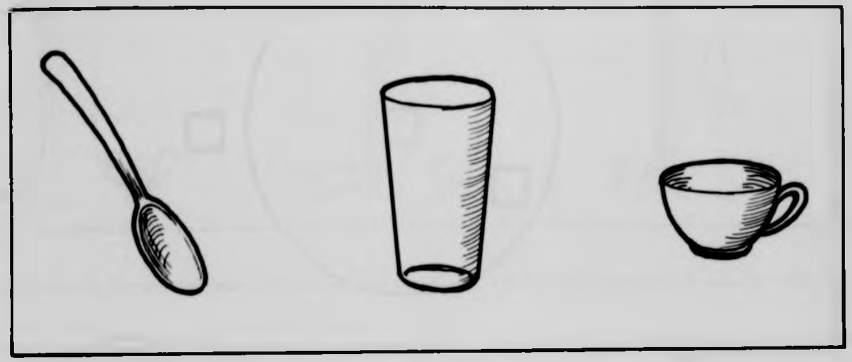

23

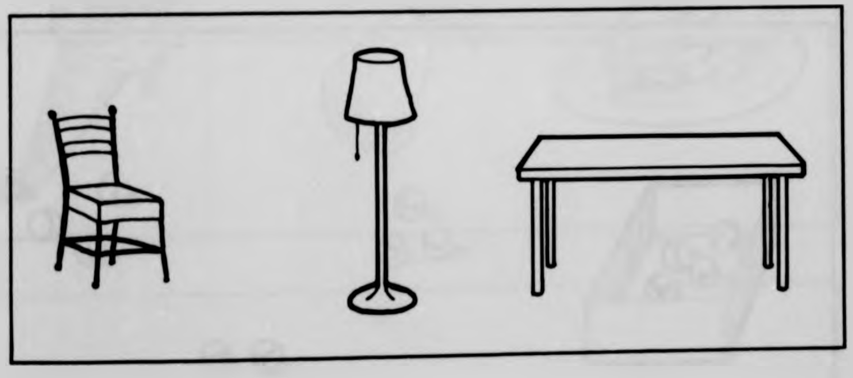

24


25

Form A
Booklet 2

NAME____





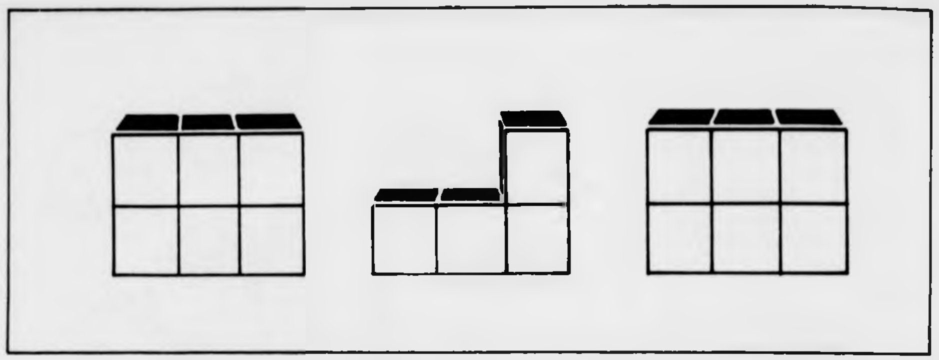
BOEHM TEST of basic concepts

Ann E. Boehm

Printed in U.S.A.

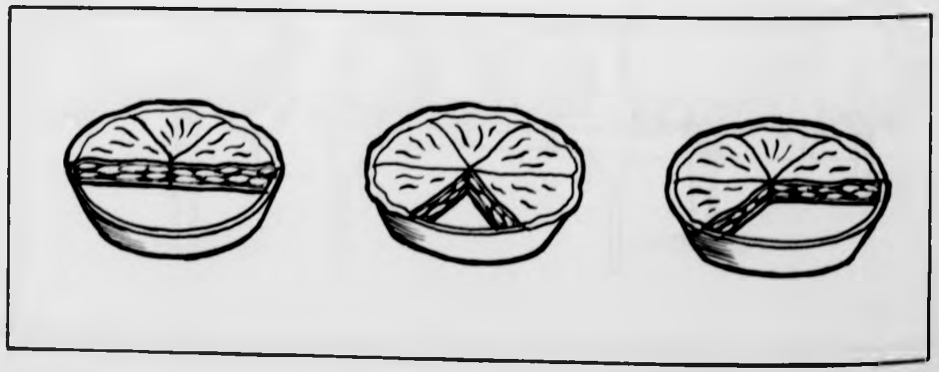
THE PSYCHOLOGICAL CORPORATION, 304 East 45th Street, New York, N.Y. 10017

70-1837


22

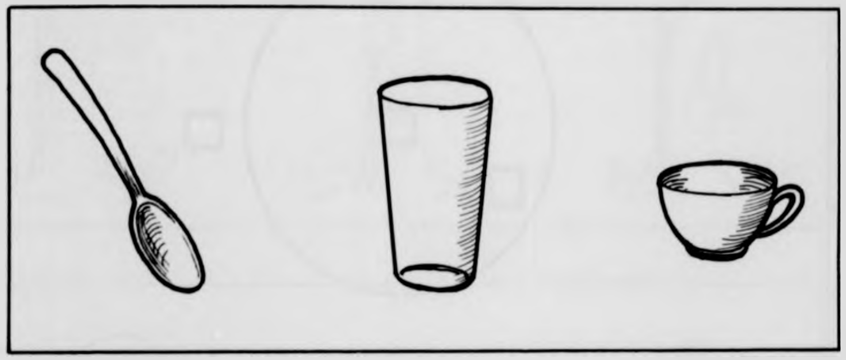

23

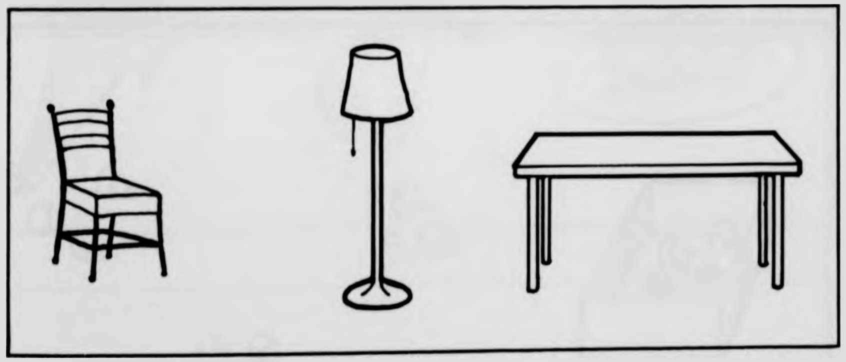

24


25

NAME ______________________________

Form A
Booklet 2





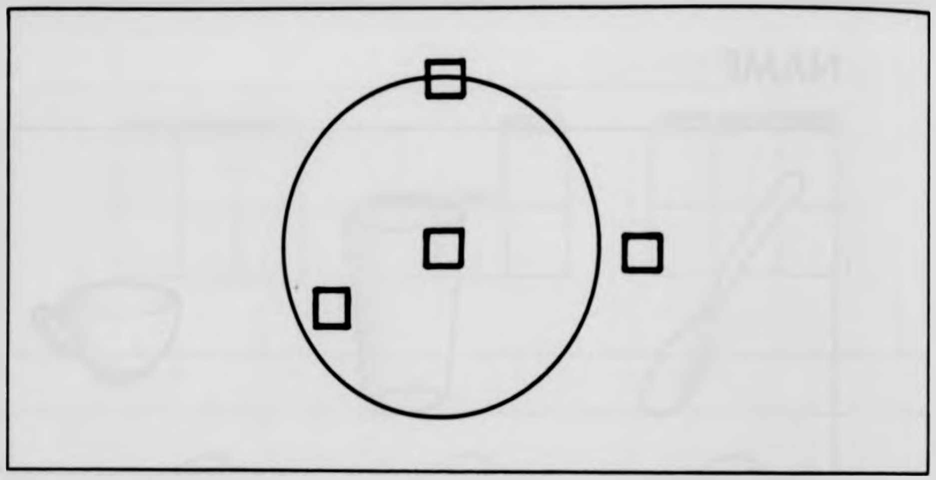
BOEHM TEST of basic concepts

Ann E. Boehm

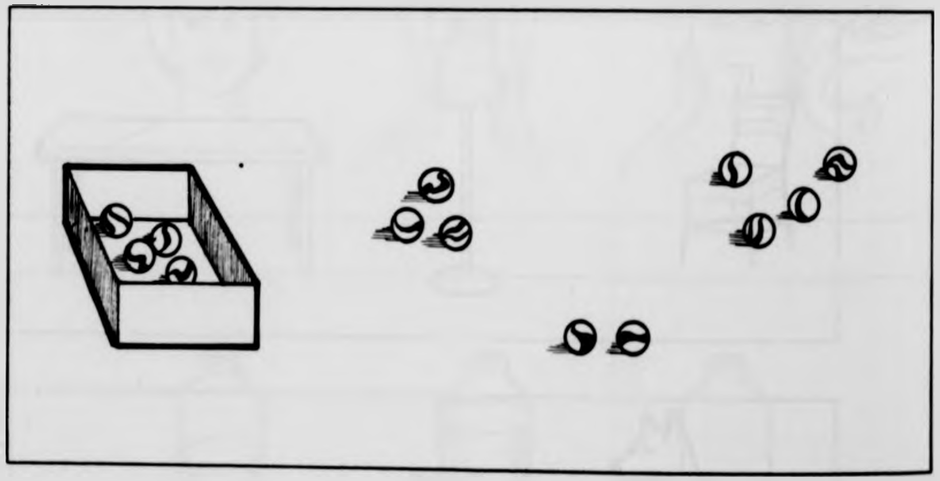
THE PSYCHOLOGICAL CORPORATION, 304 East 45th Street, New York, N.Y. 10017

Printed in U.S.A.

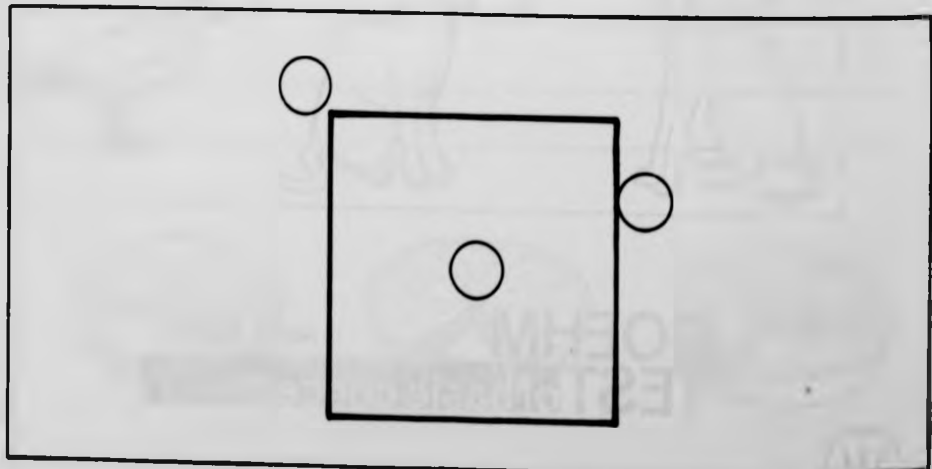
70-1837



26

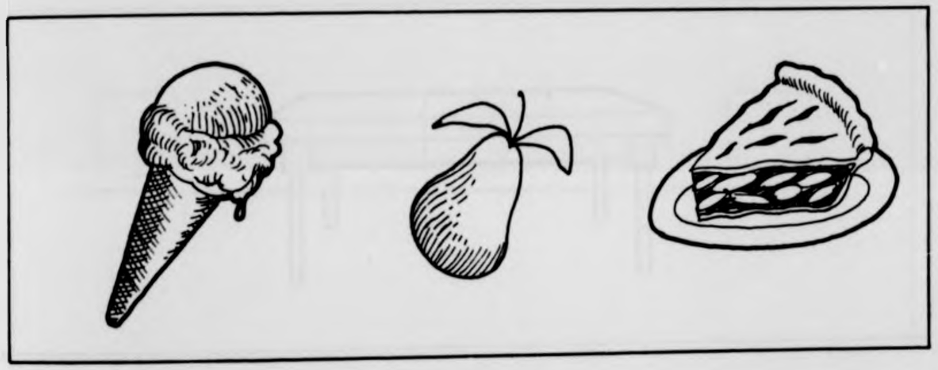


27

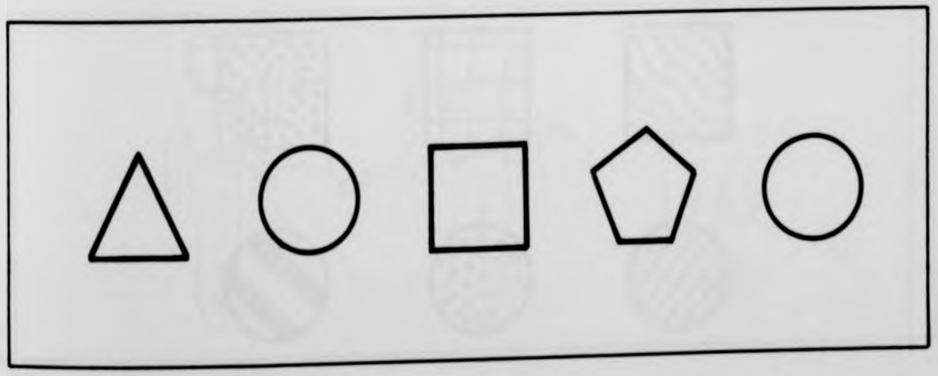


28

29

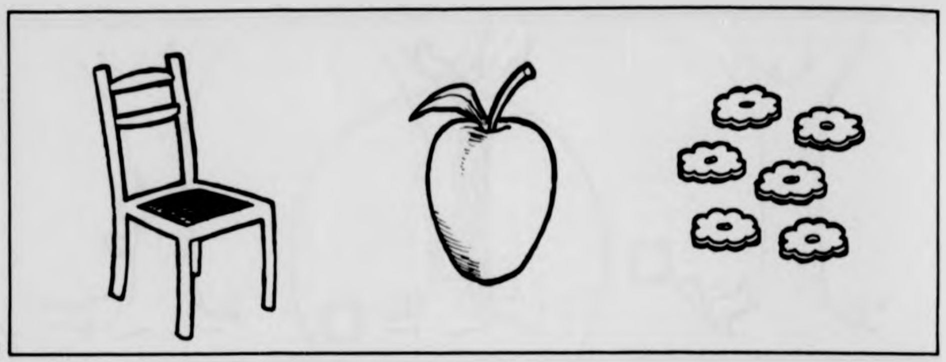


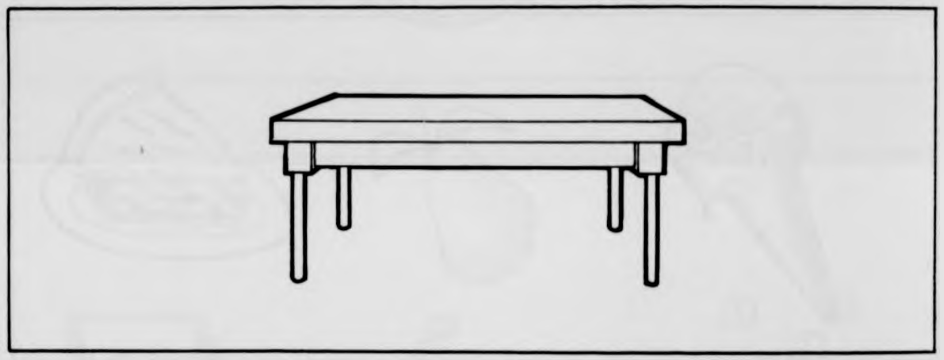
30

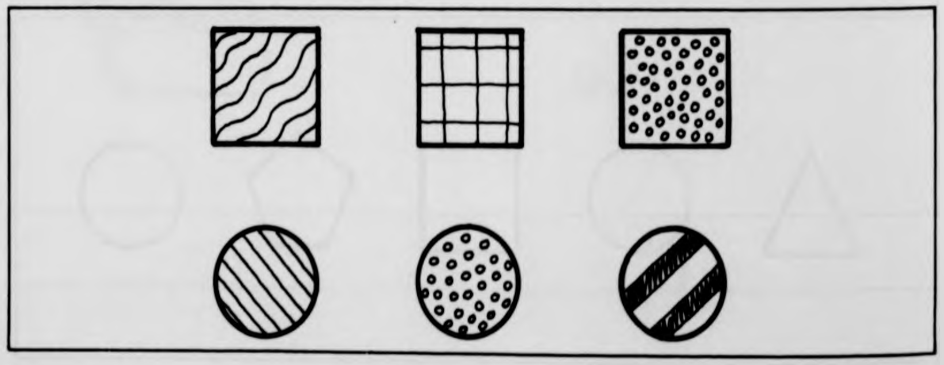


31

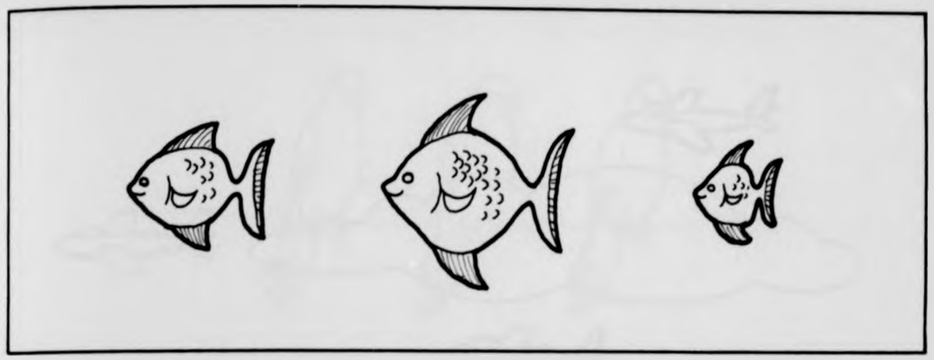
32

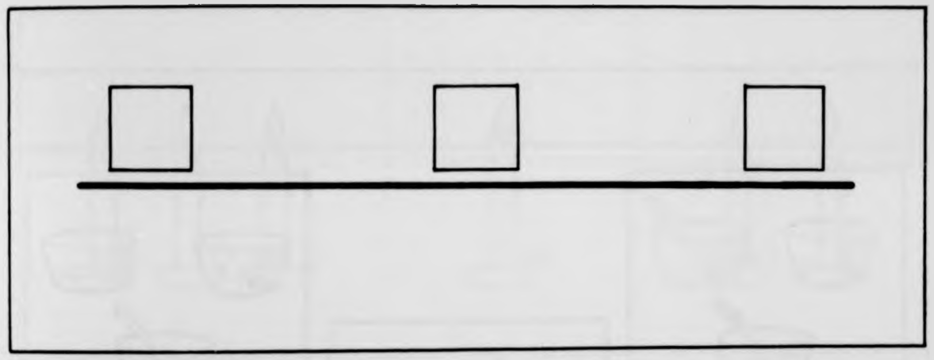

33


34

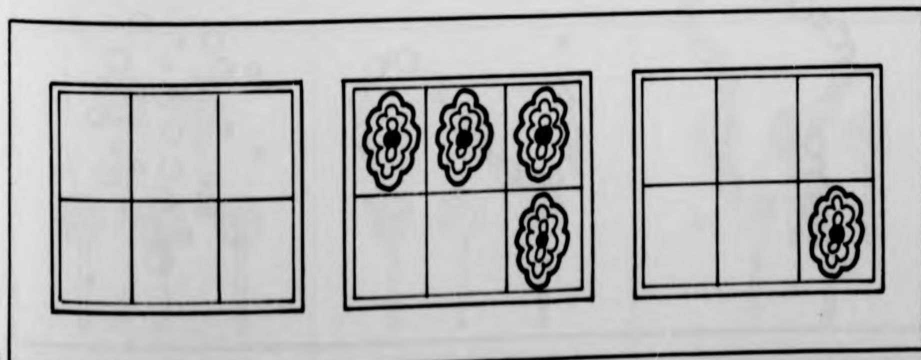

35

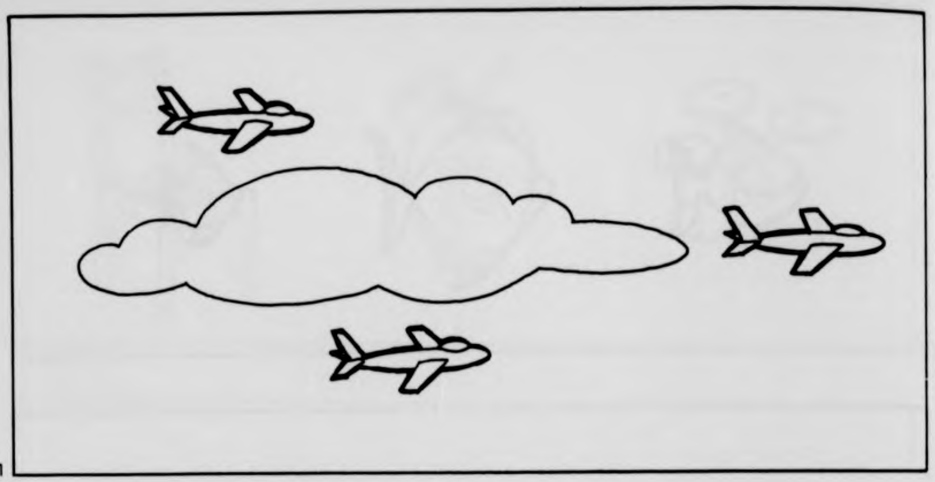

36

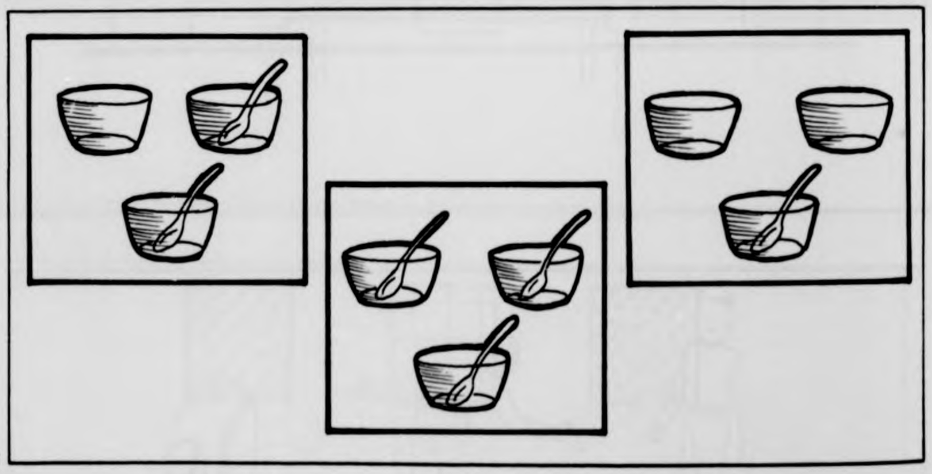

37

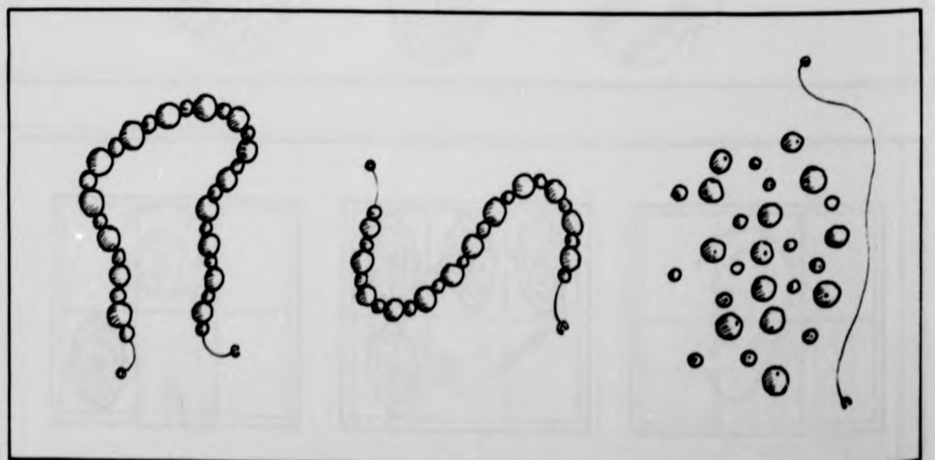

38

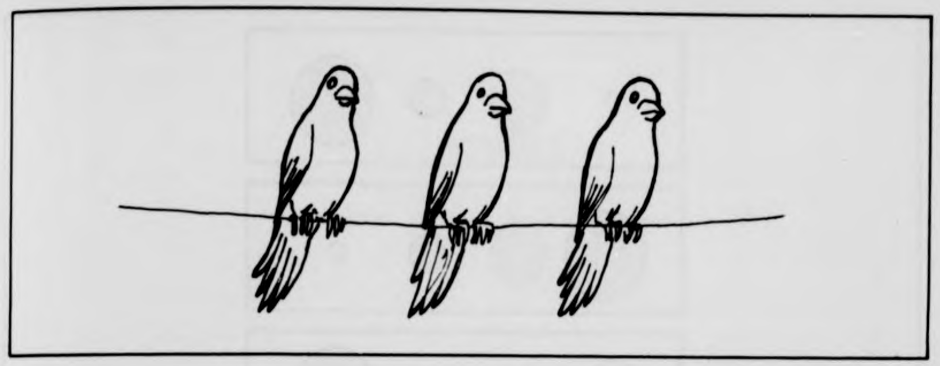

39

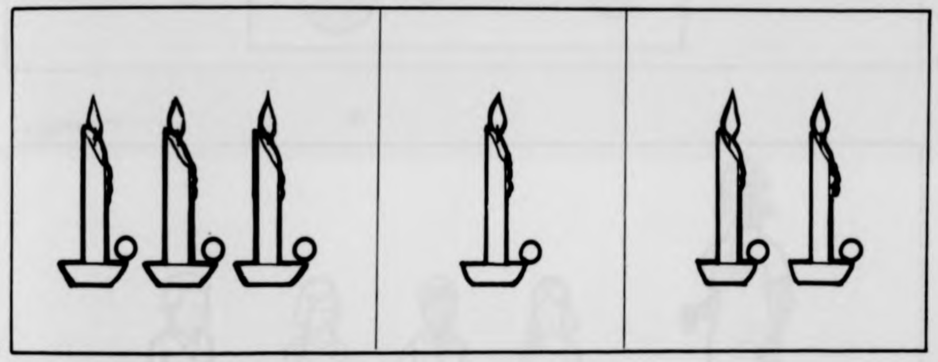

40

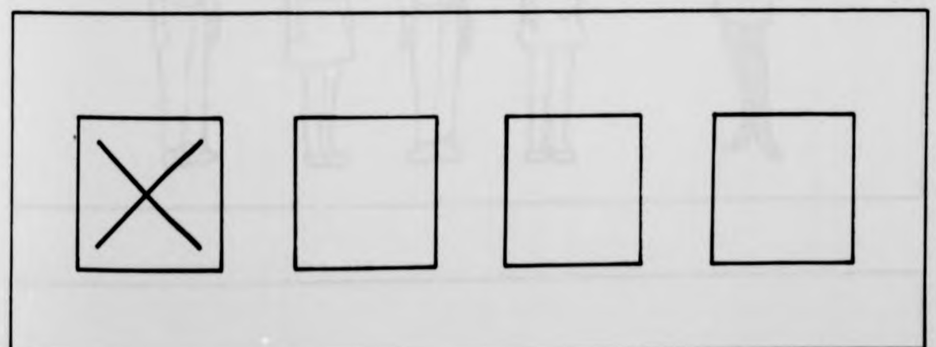

41

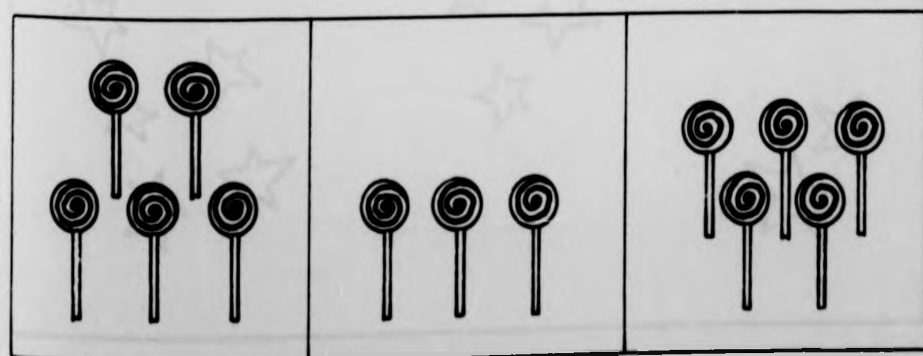

42

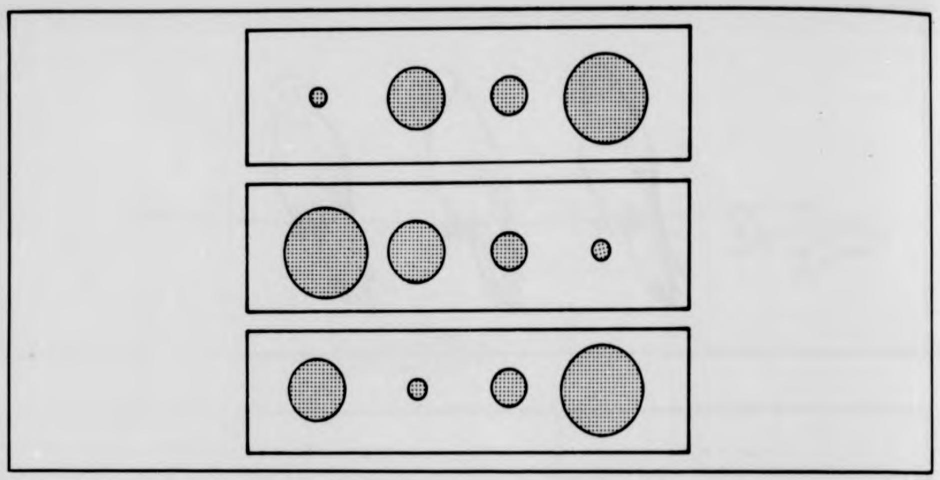

43


44

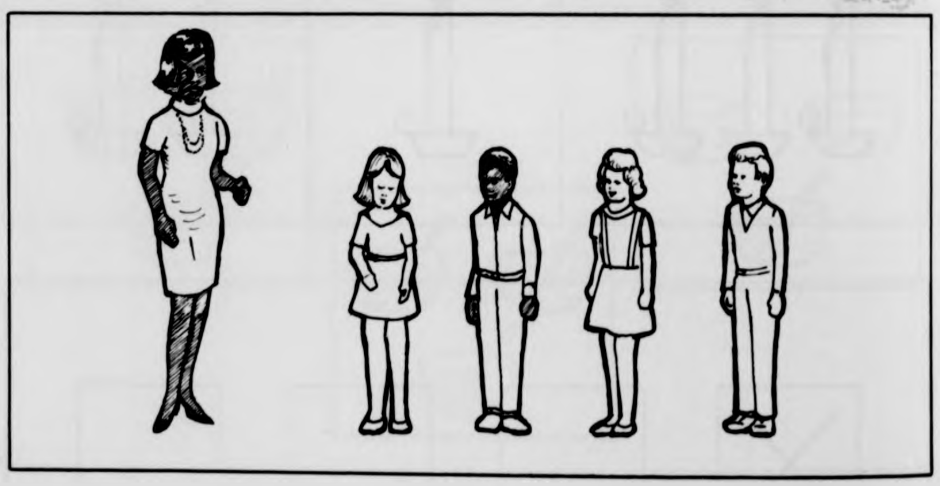

45


46

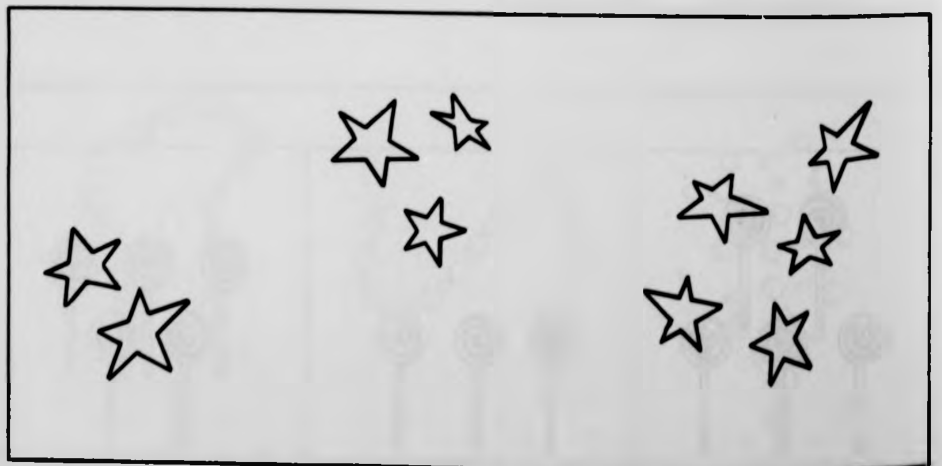

47



48



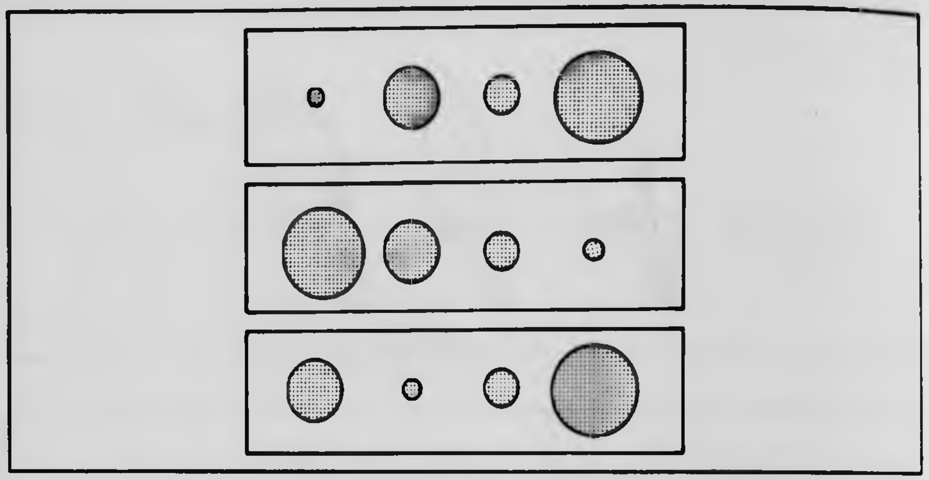
49



50

Appendix 3

The BTBC Form B, Booklets 1 and 2.



48

49

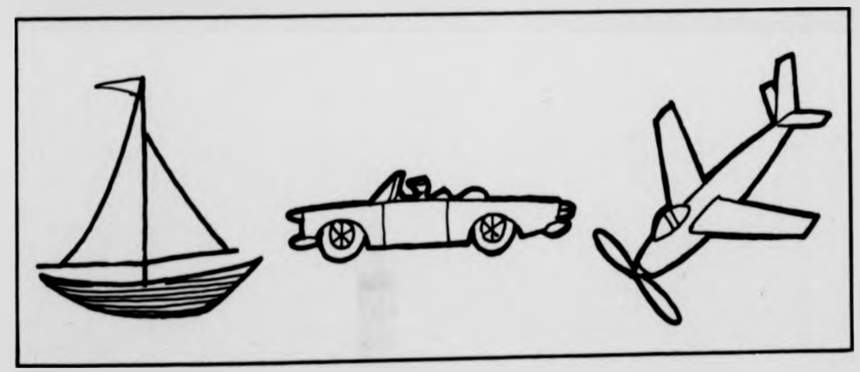
50

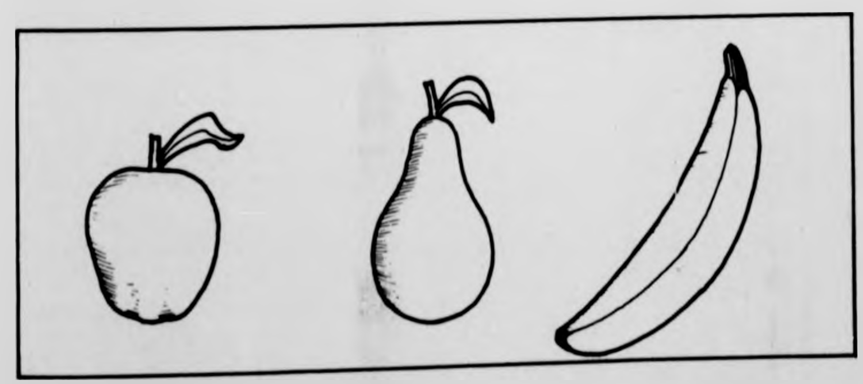
Appendix 3

The BTBC Form B, Booklets 1 and 2.

Form **B**
Booklet **1**

NAME ______________________________





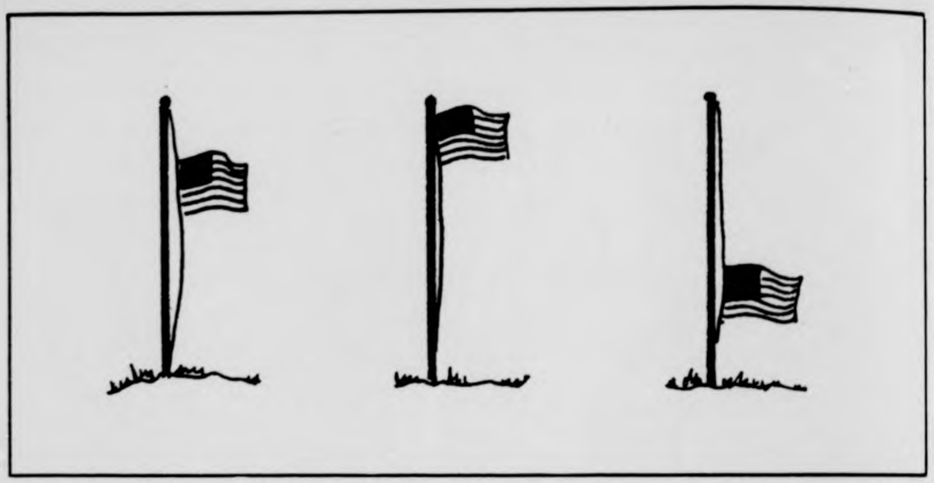
BOEHM TEST of basic concepts

Ann E. Boehm

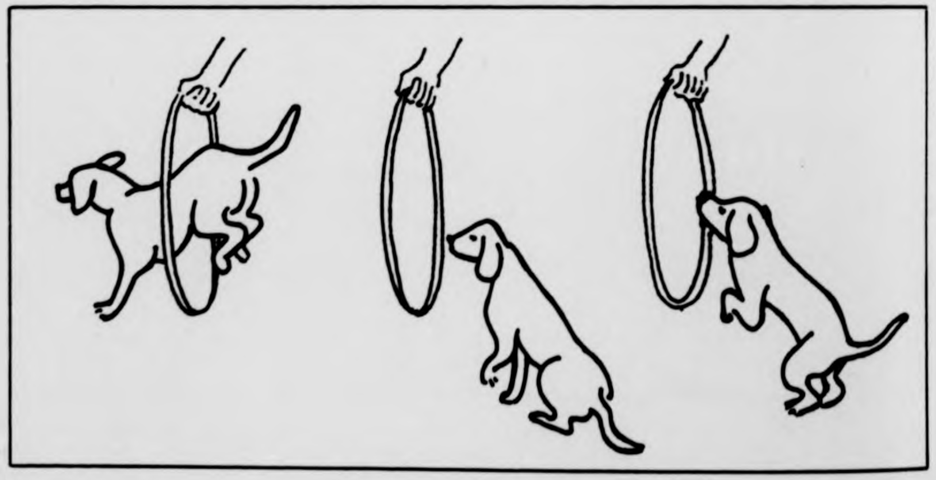
Printed in U.S.A.

THE PSYCHOLOGICAL CORPORATION, New York

76-2011 9-010822

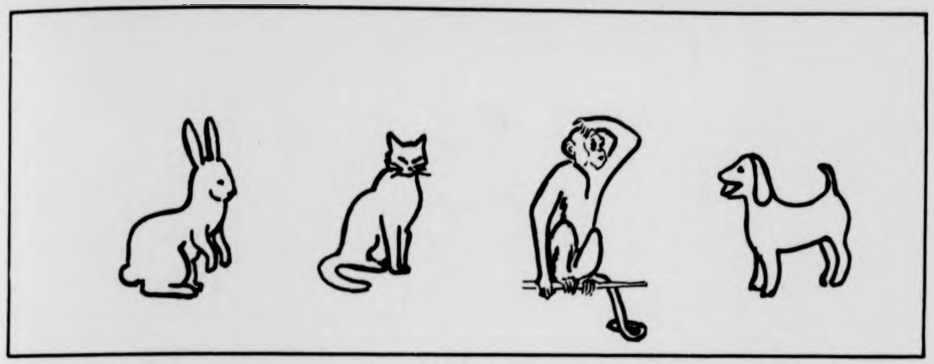


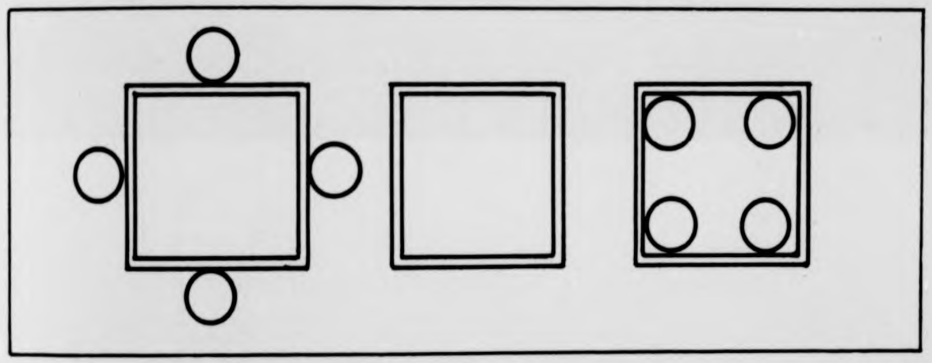
1

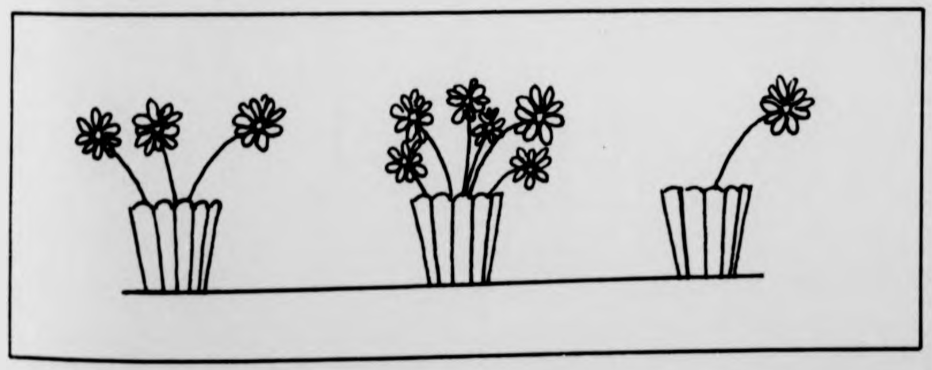


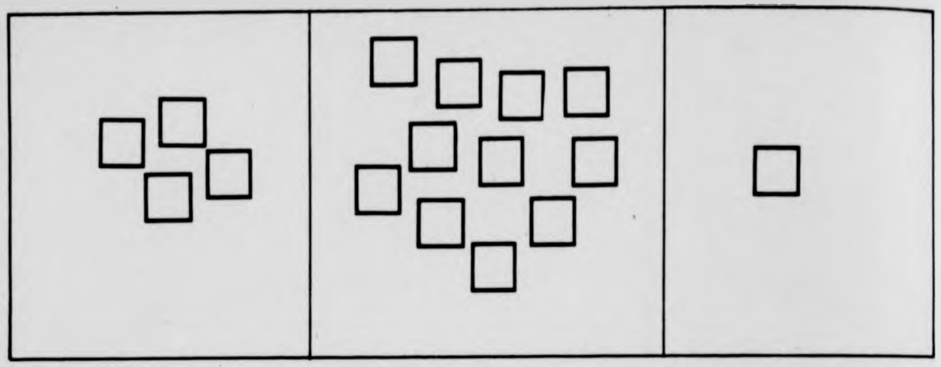
2

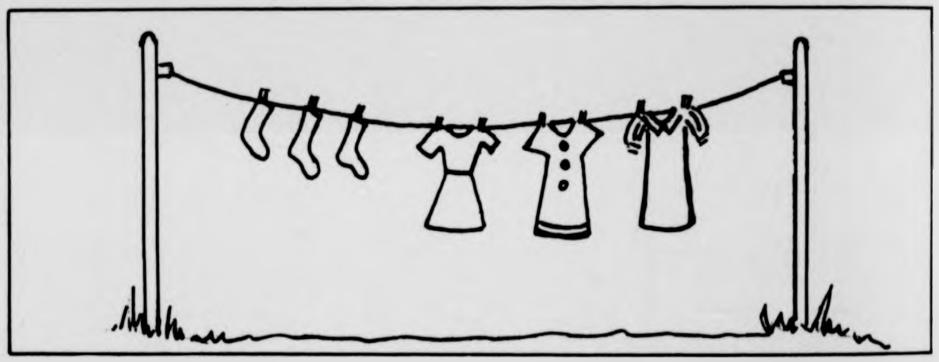
3

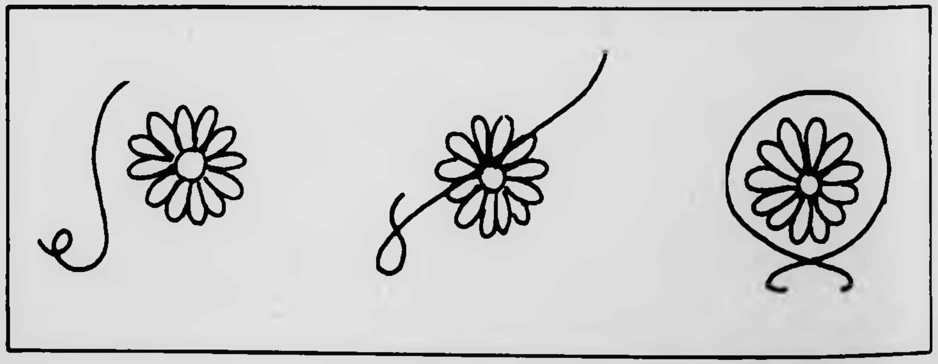


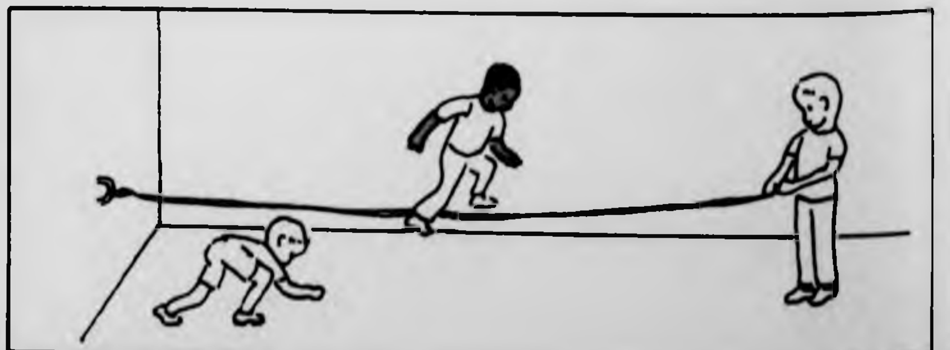


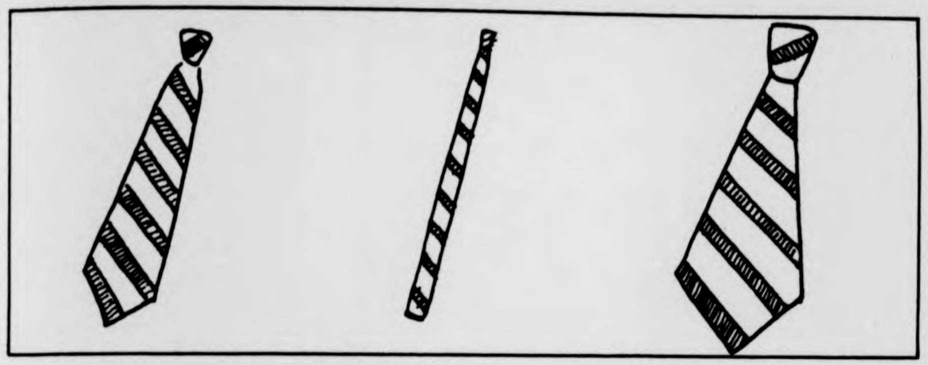


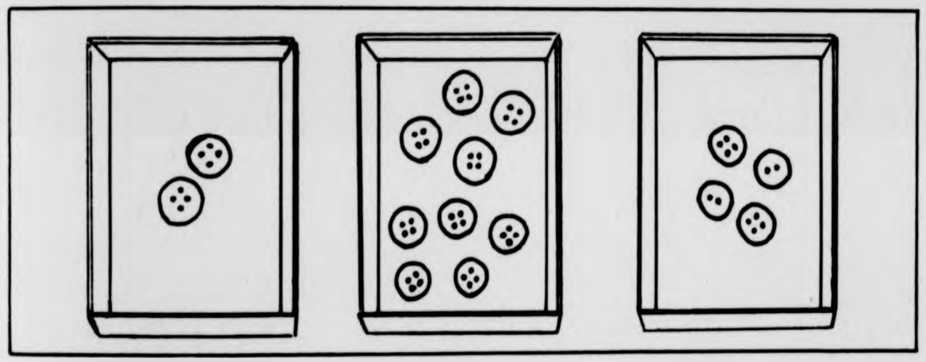

8

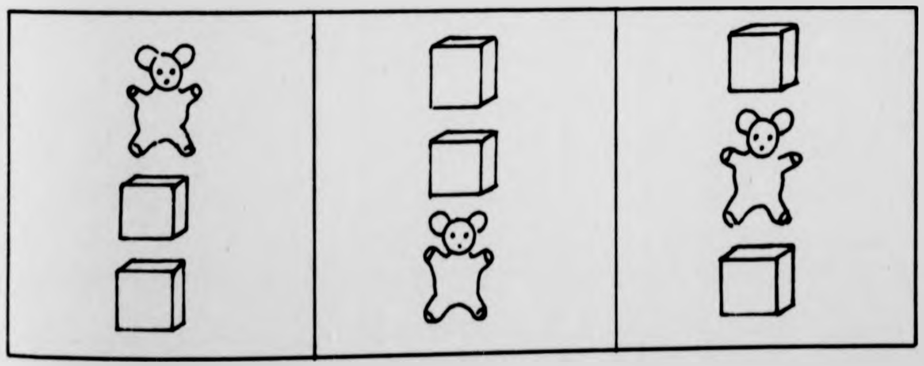

9


10

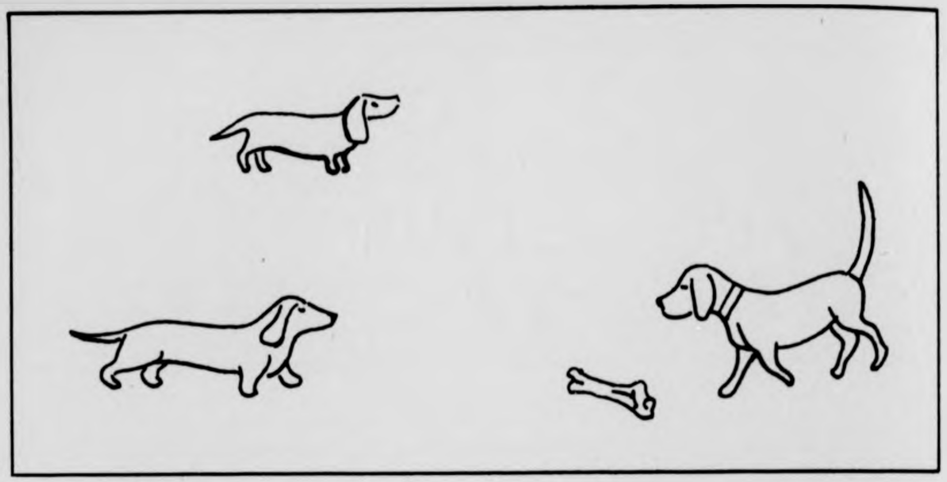

11


12

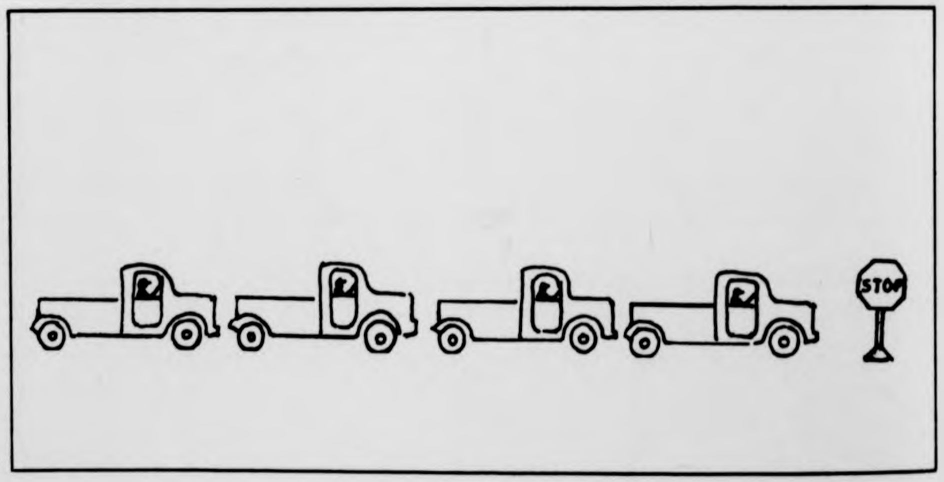

13


14

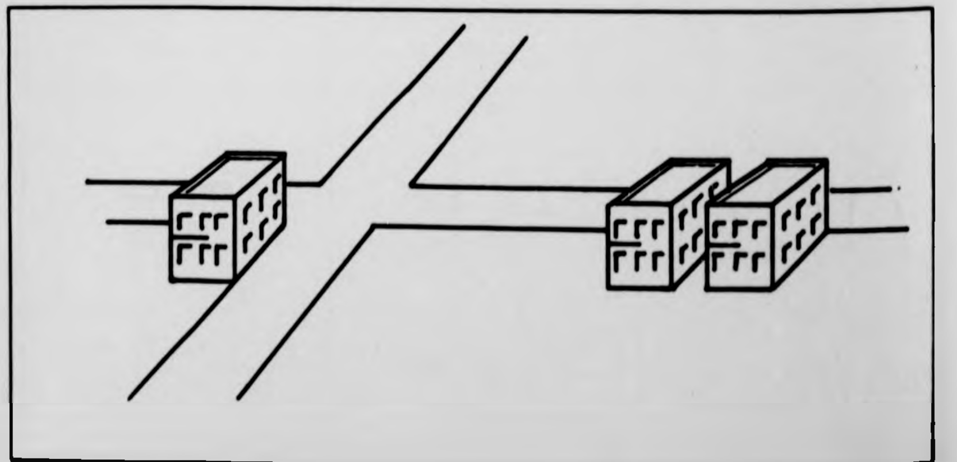

15



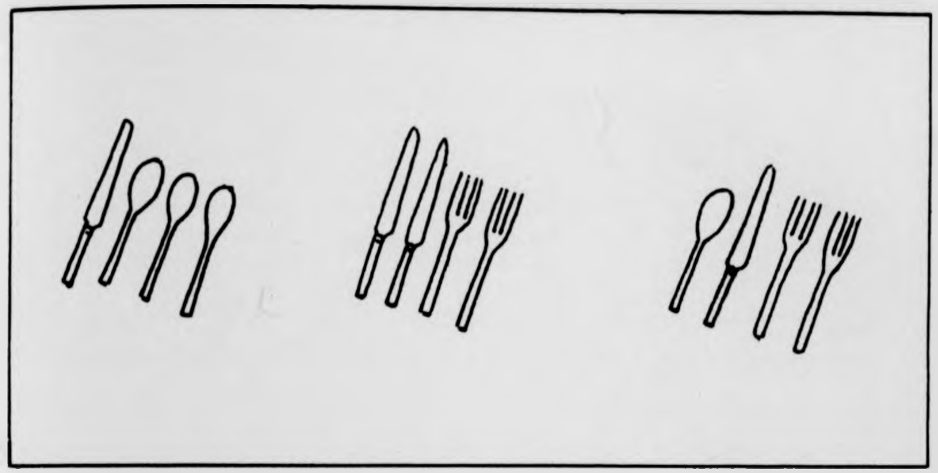
16



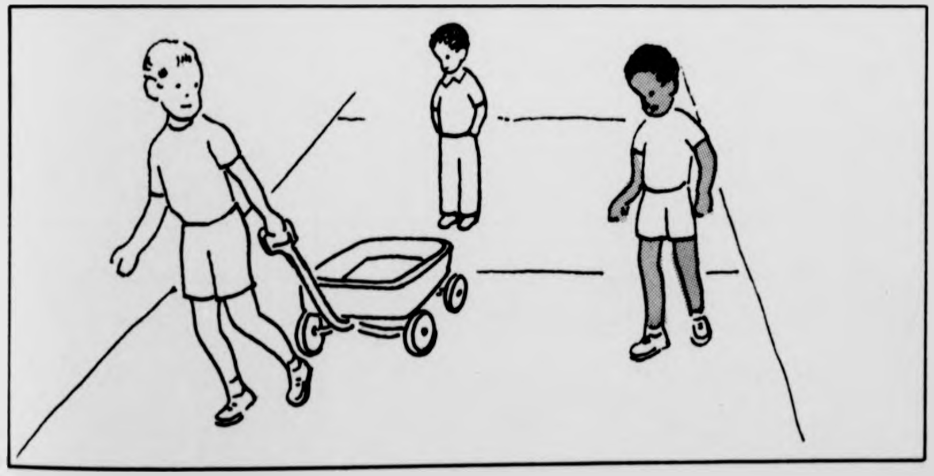
17



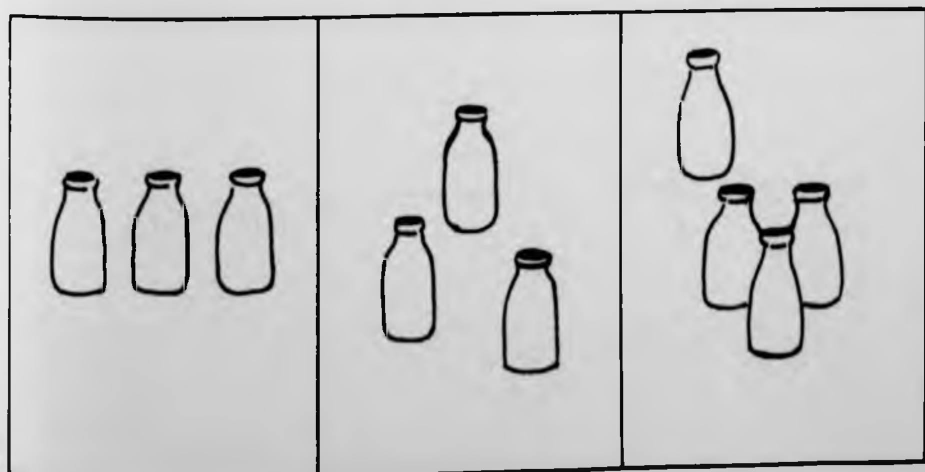
18



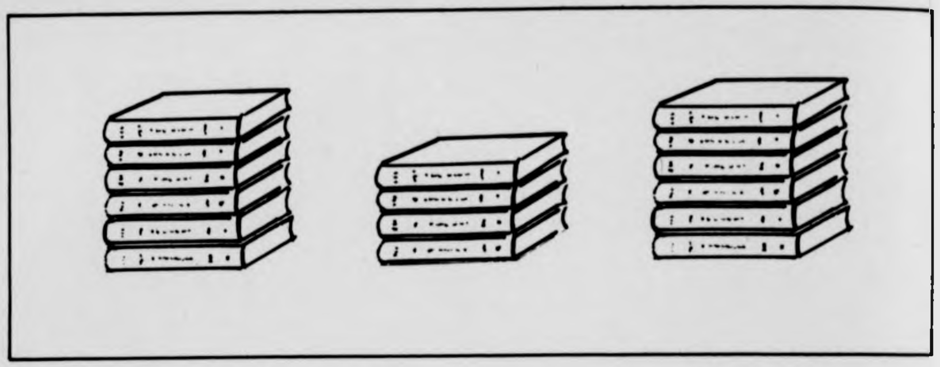
19



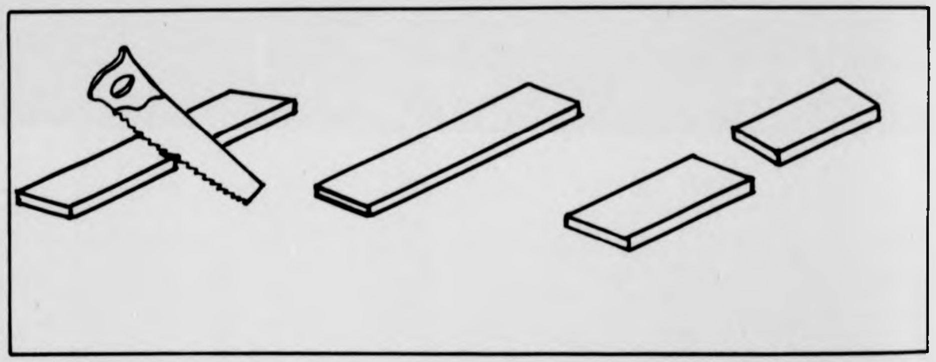
20



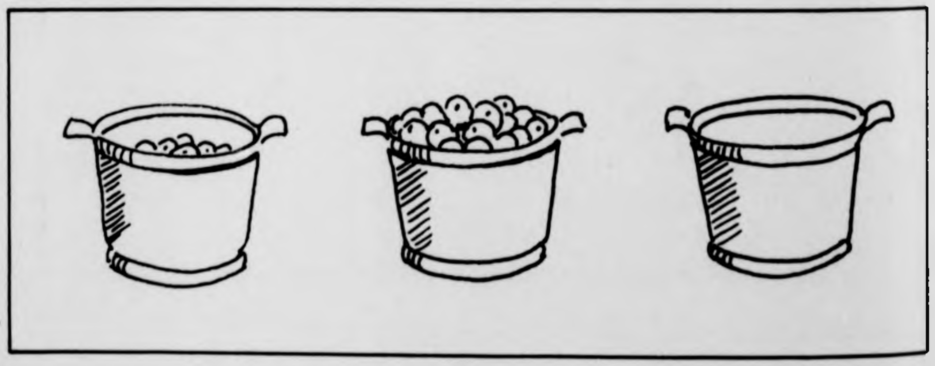
21



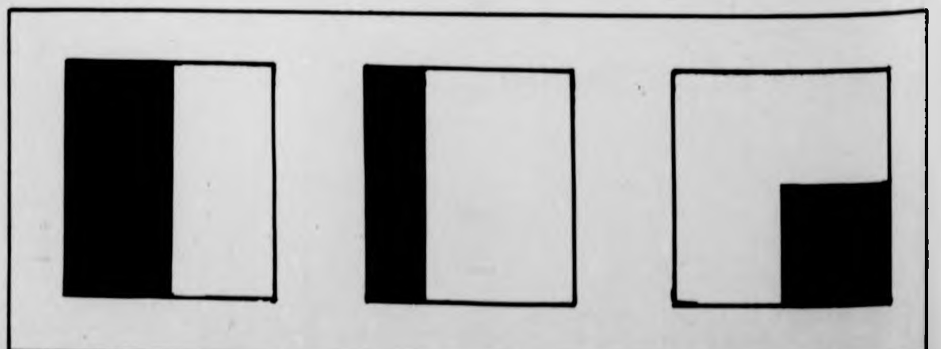
22



23



24

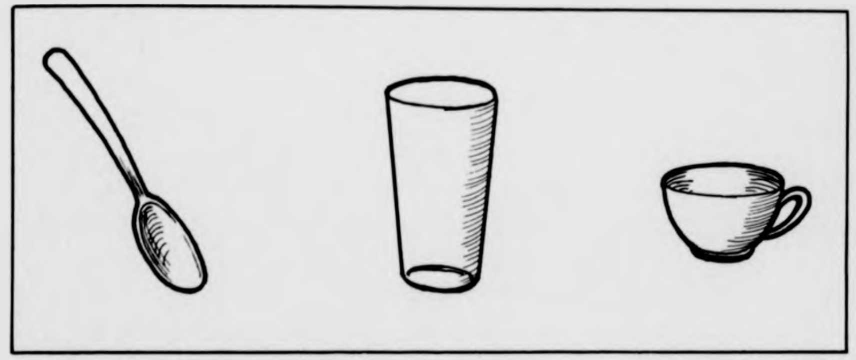


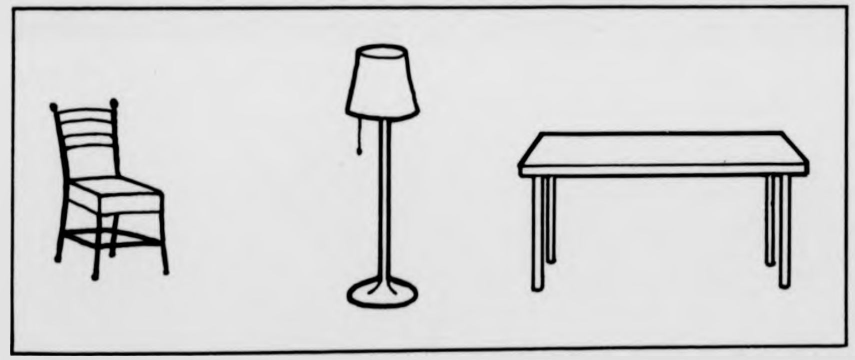
25

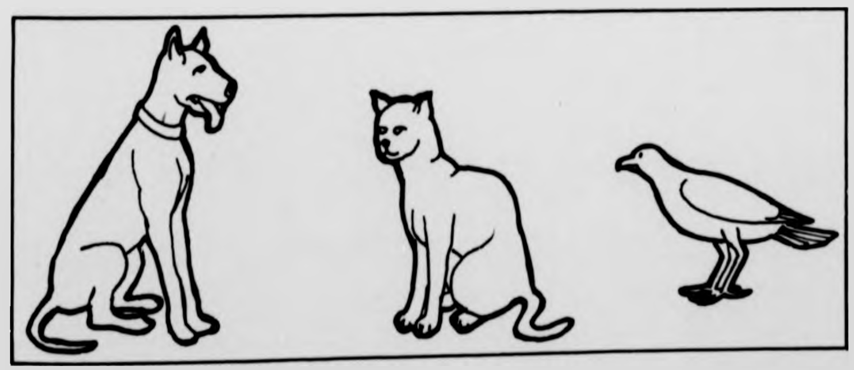
NAME ____________________________________

Form **B**

Booklet **2**



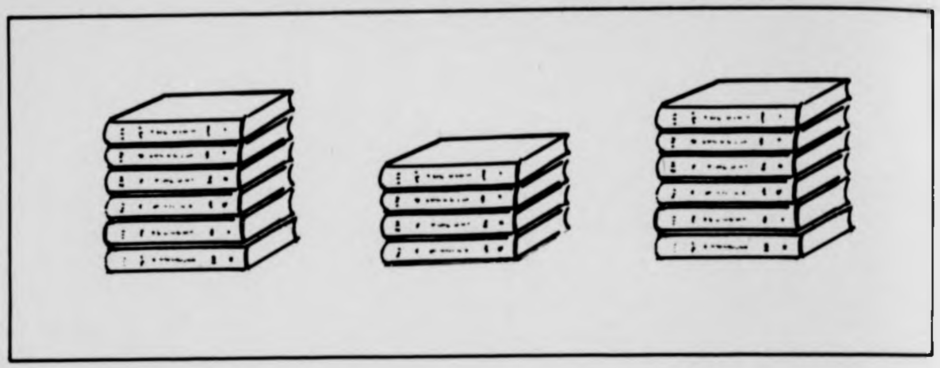




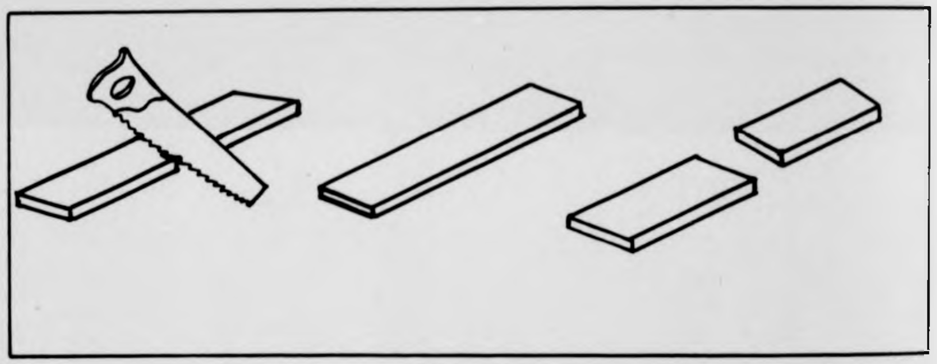
BOEHM TEST of basic concepts

Ann E. Boehm

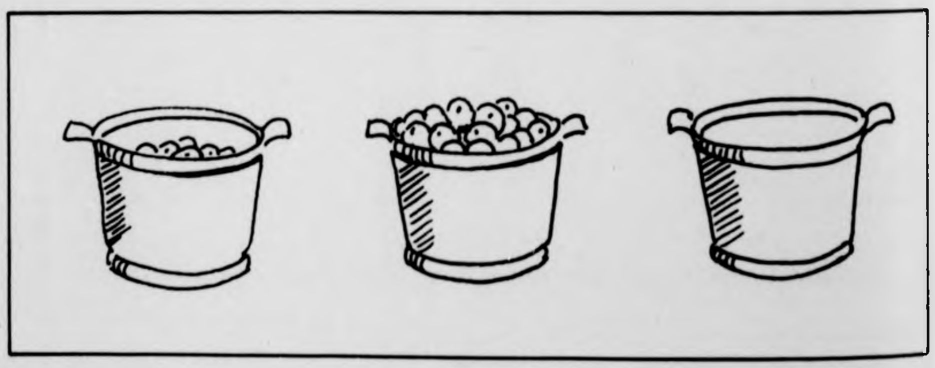
Printed in U.S.A.



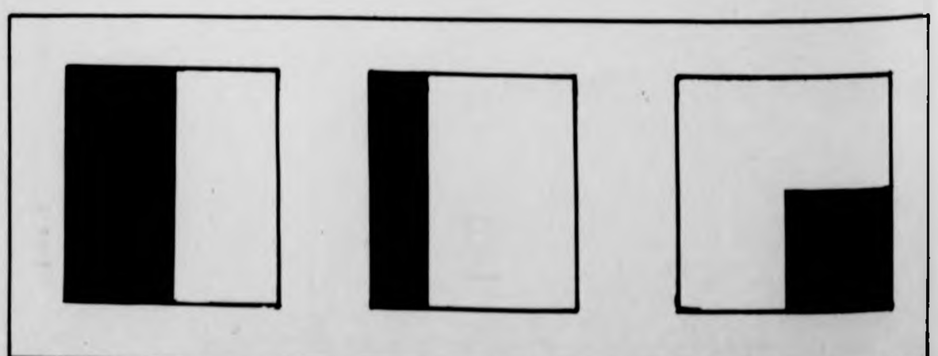
22



23



24

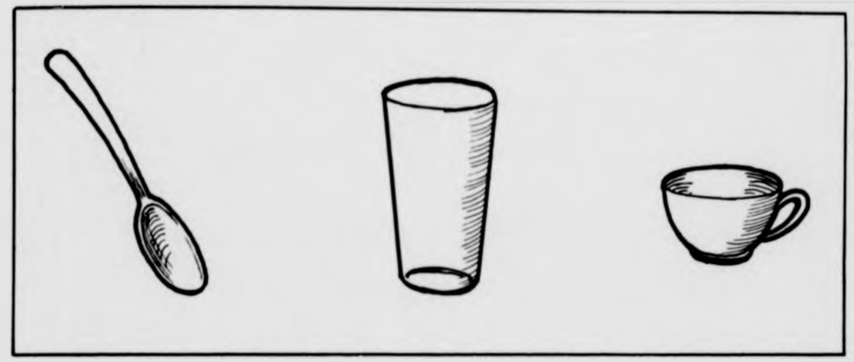


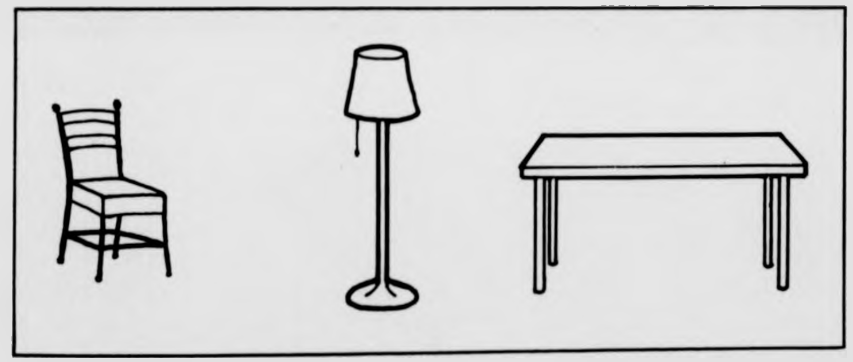
25

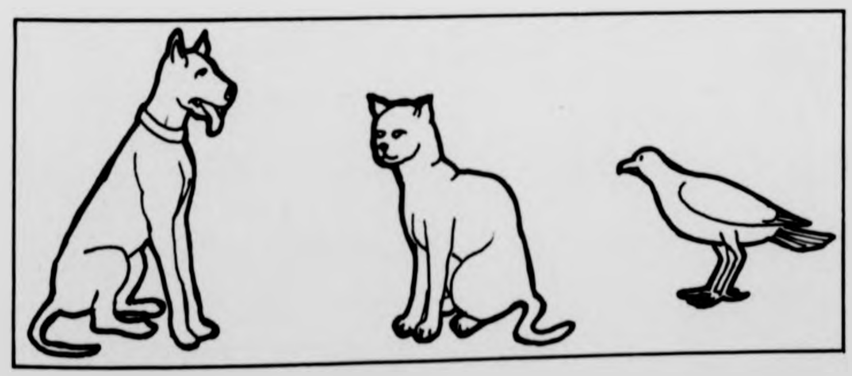
NAME________________________________

Form **B**

Booklet **2**





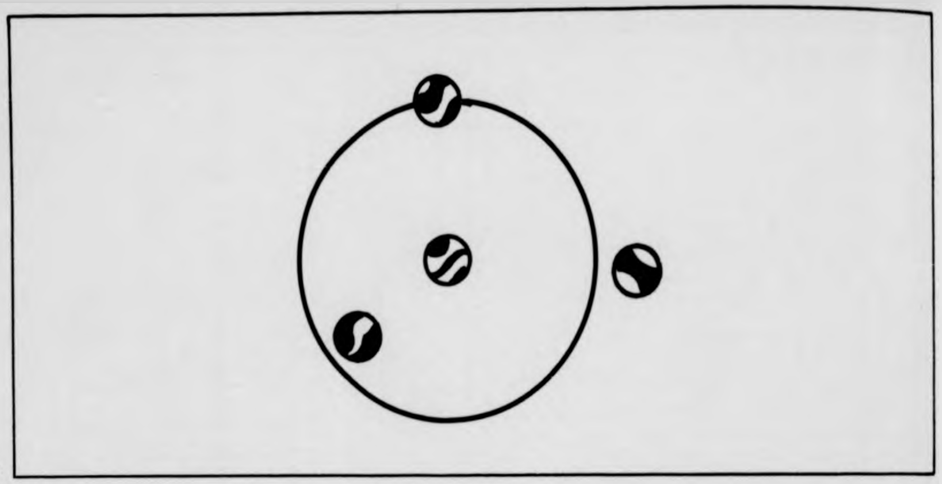


BOEHM TEST of basic concepts

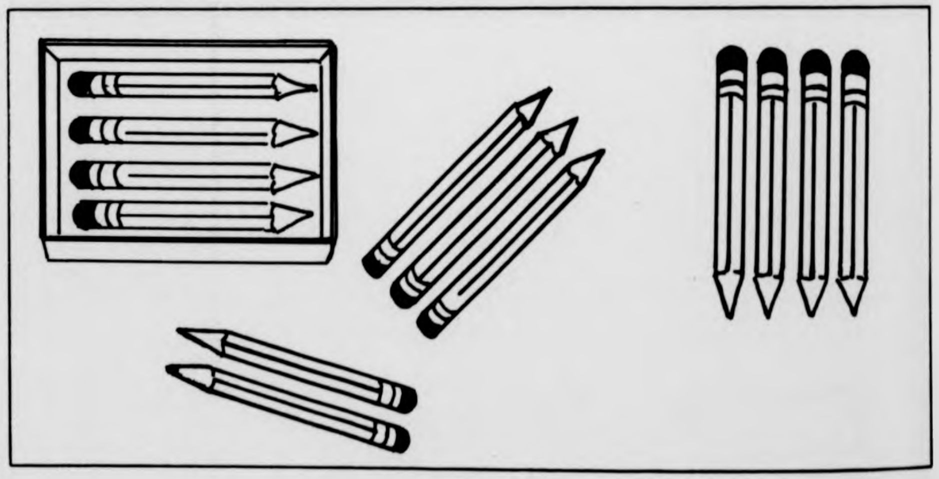
Ann E. Boehm

Printed in U.S.A.

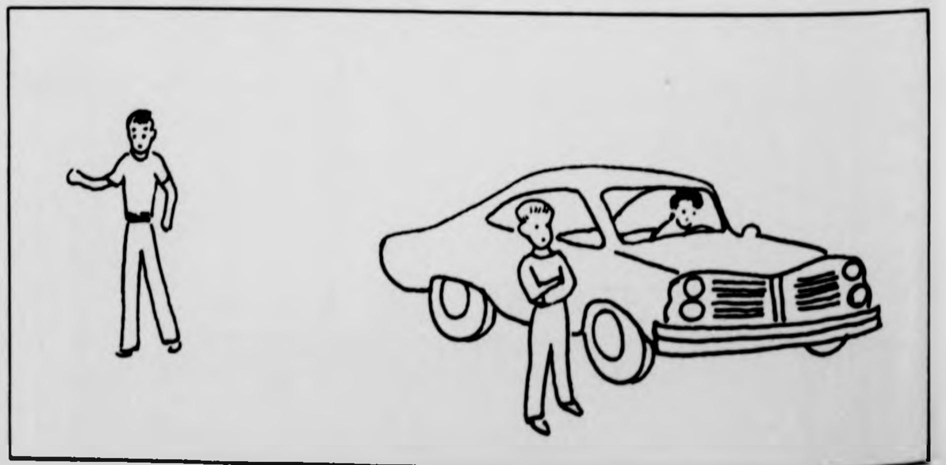
THE PSYCHOLOGICAL CORPORATION, New York



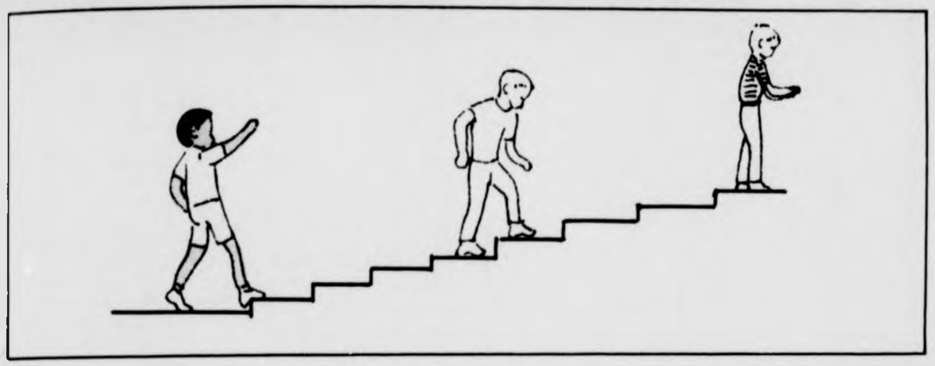
26



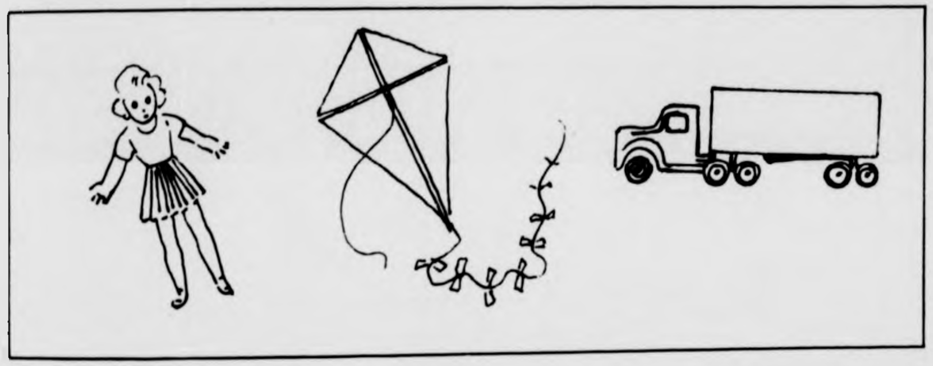
27



28

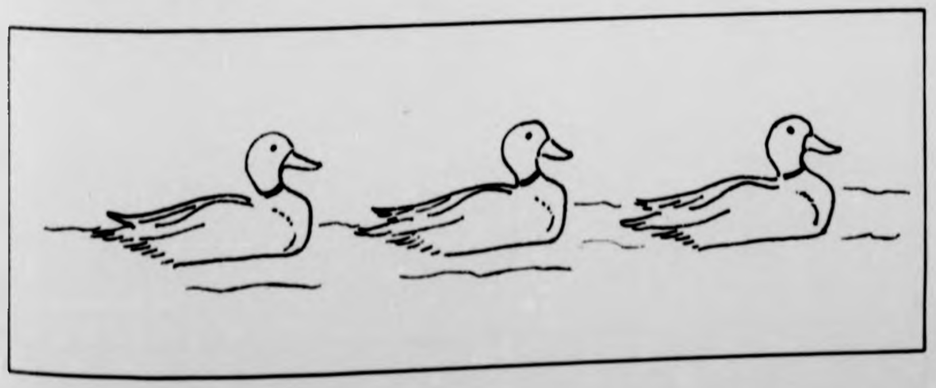


29

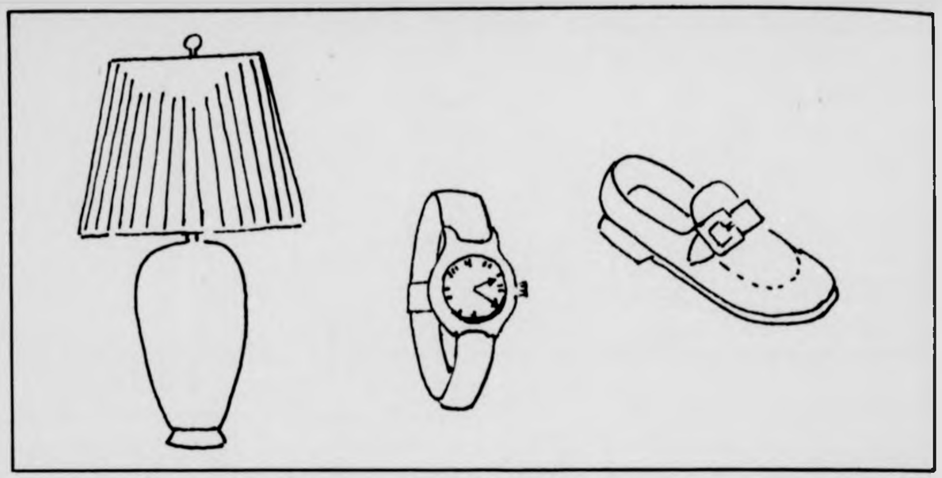


30

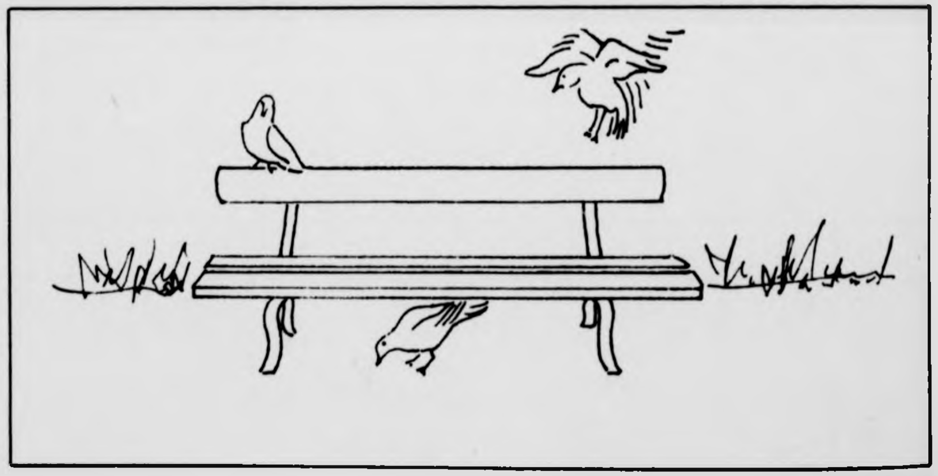
31



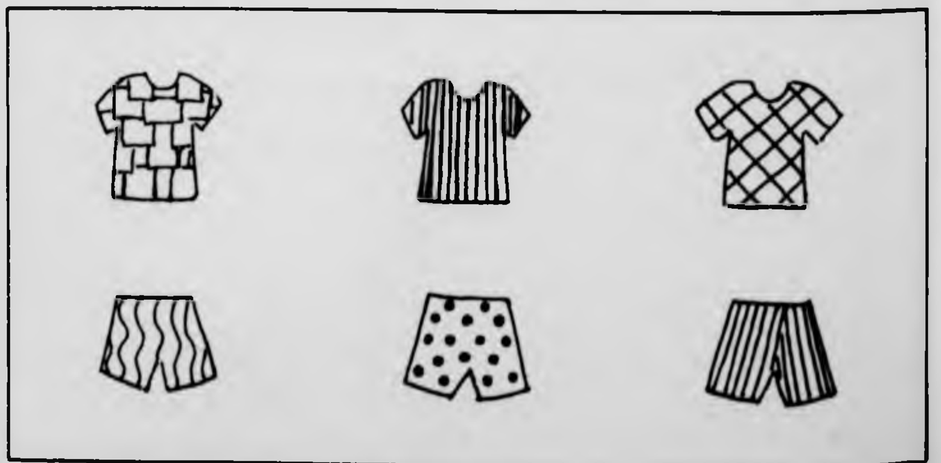
32



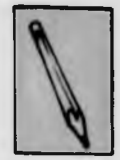
33

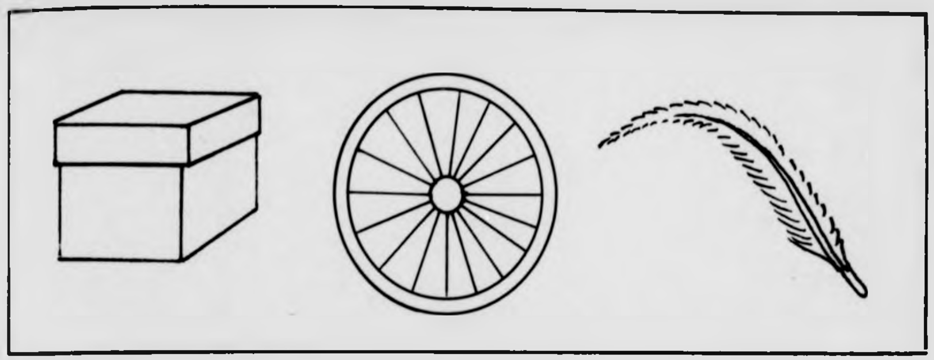


34



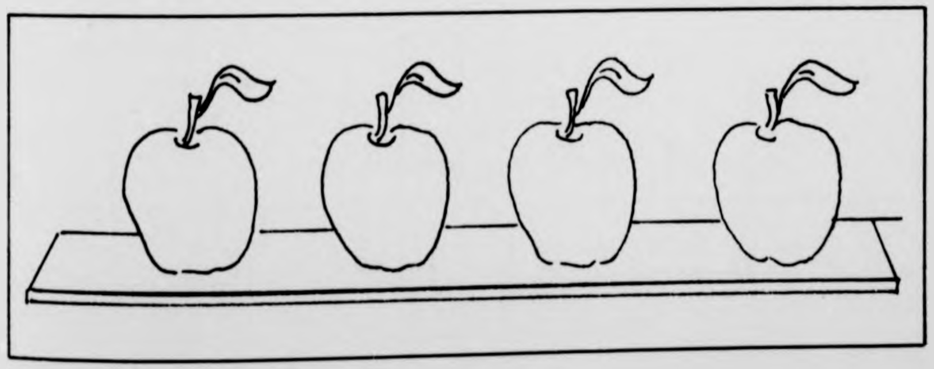
35



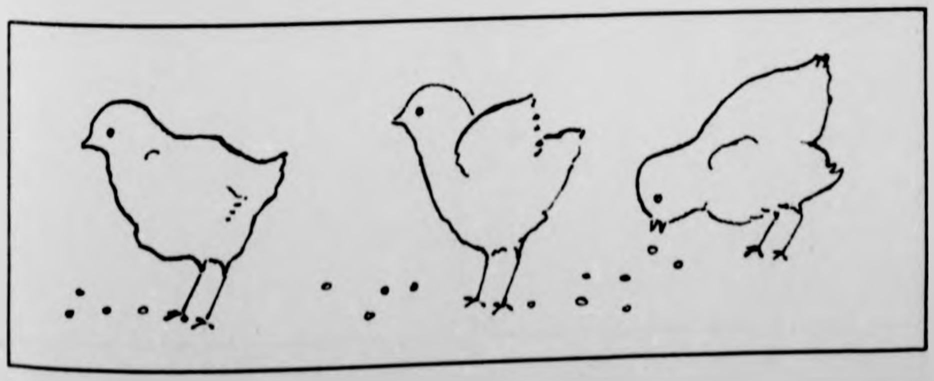


36

37

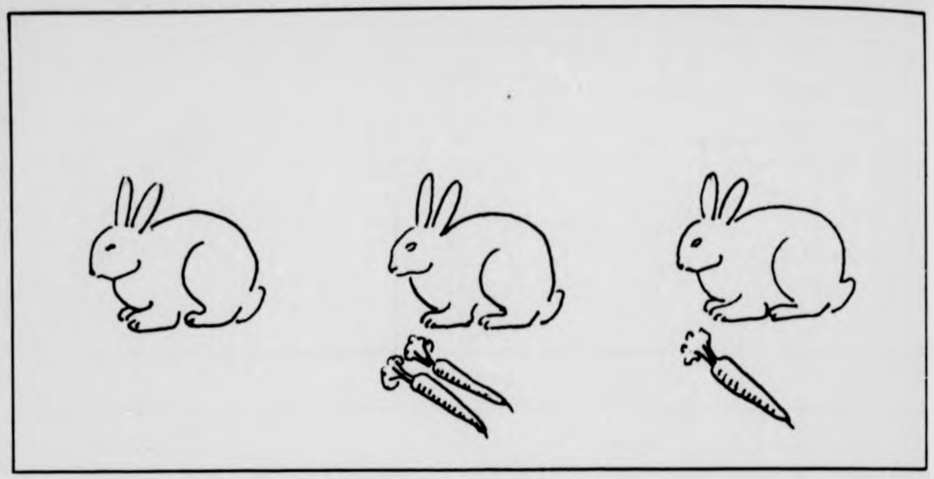


38

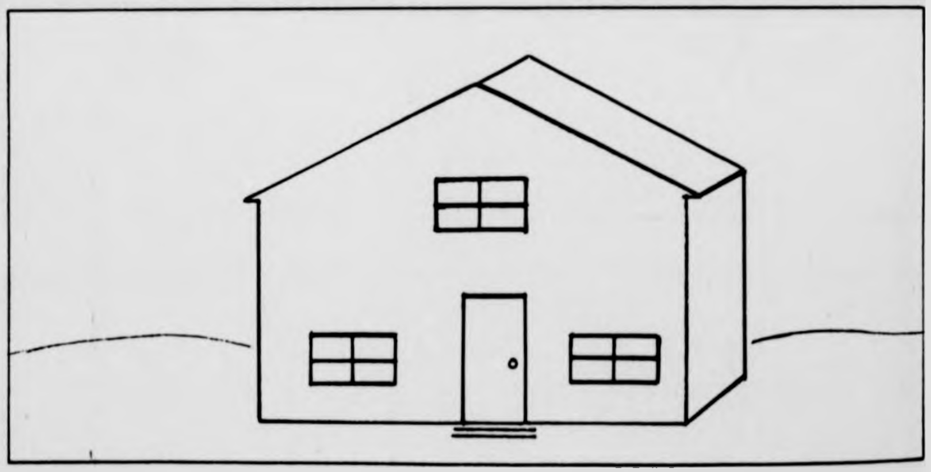


39

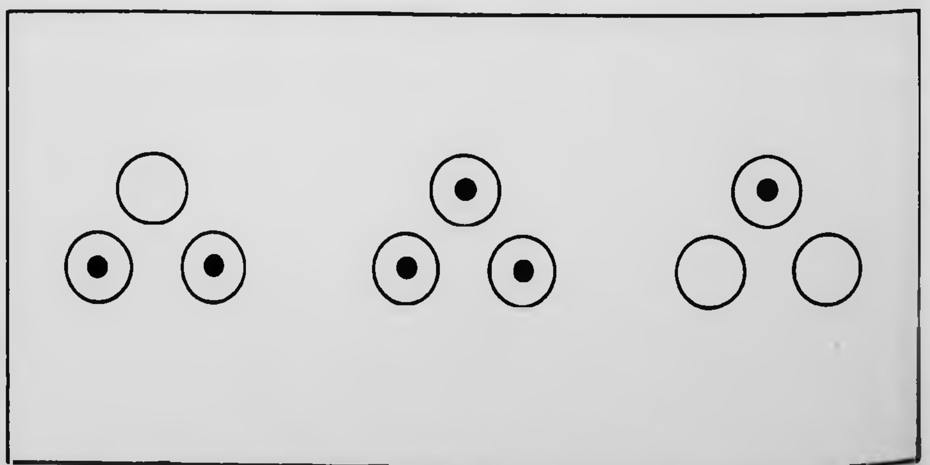
40



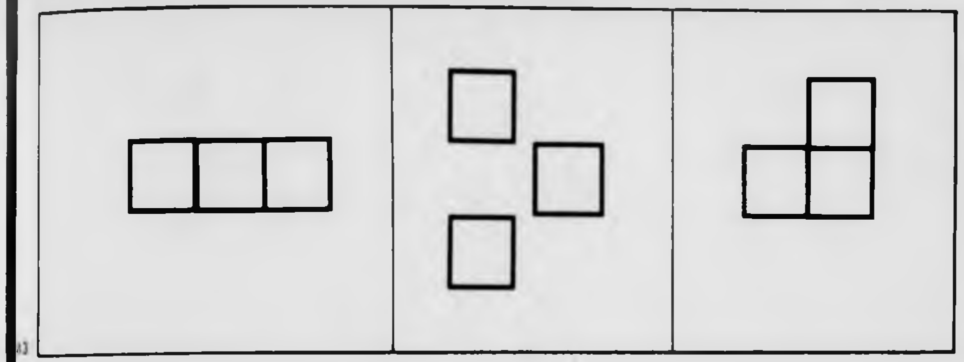
41



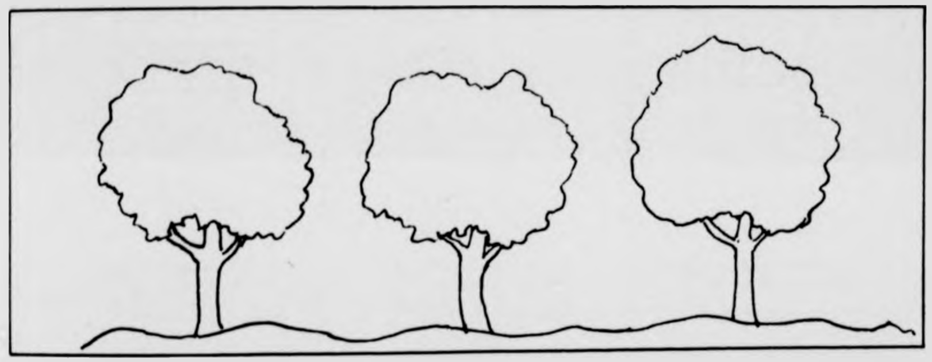
42



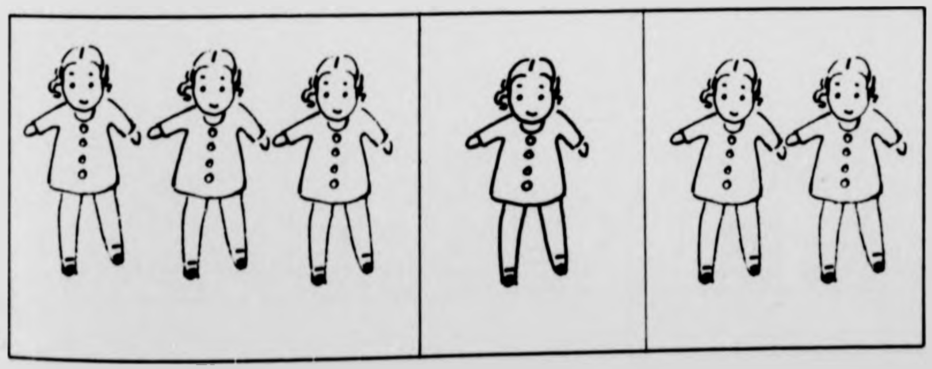
43



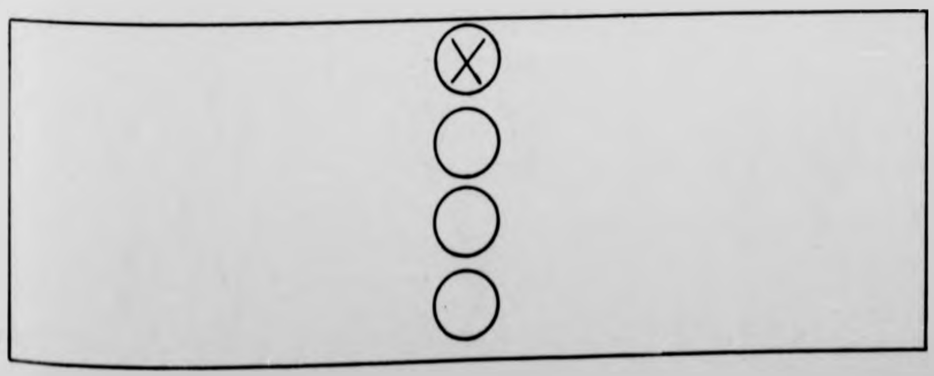
44



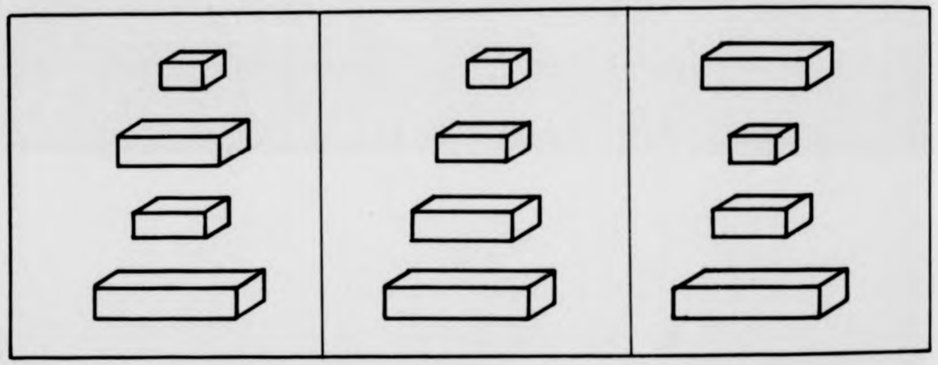
5



6



47



48

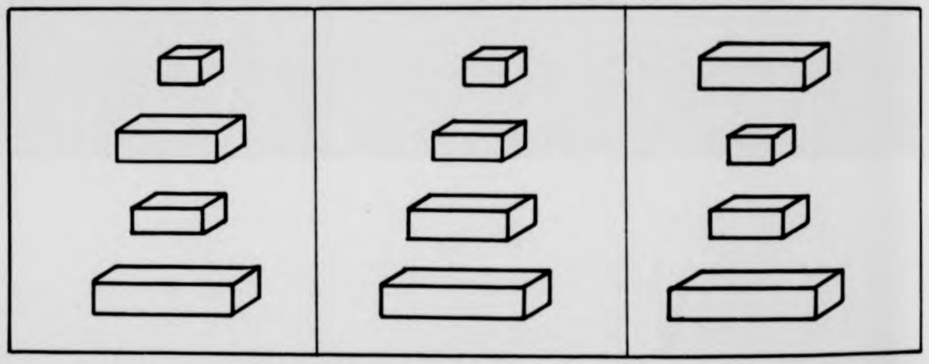
49

50

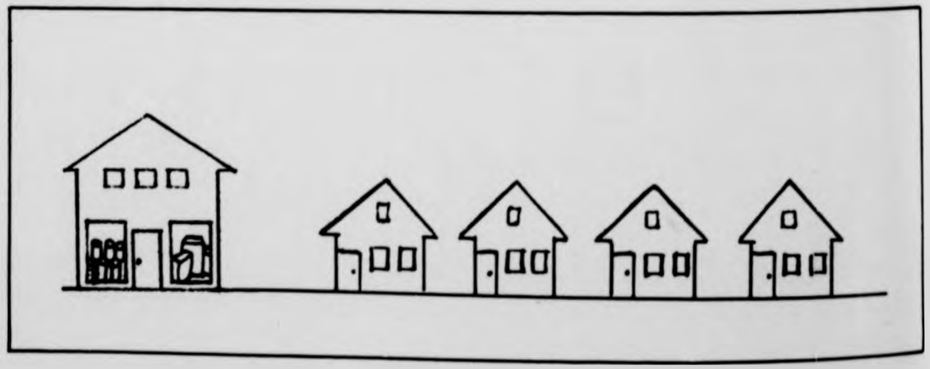
Appendix 4

The Boehm Test of Basic Concepts Manual, 1971 Edition, Forms A and B.

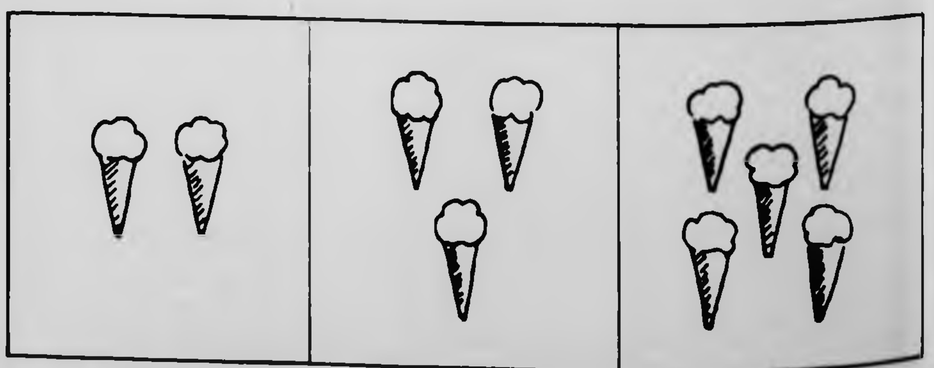
47



48



49



50

Appendix 4

The Boehm Test of Basic Concepts Manual, 1971 Edition, Forms A and B.

BOEHM TEST of basic concepts MANUAL

Ann E. Boehr

1971 Edition
Forms A and

The Psychological Corporation

TABLE OF CONTENTS

LIST OF TABLES

THE PSYCHOLOGICAL CORPORATION, 304 East 45th Street, New York, N. Y. 10017

Printed in U.S.A.

74-124M

INFORMATION FOR THE TEACHER

General Description

The *Boehm Test of Basic Concepts* (*BTBC*) is designed to measure children's mastery of concepts considered necessary for achievement in the first years of school. The test is read aloud by the teacher, and is appropriate for use with children in the three grades—kindergarten, 1, and 2. *BTBC* results may be used both to identify *children* with deficiencies in this area, and to identify individual *concepts* on which the children could profit from instruction.

Two alternate forms of the *BTBC* are now available. Form A was published in 1969; Form B, which yields a total raw score equivalent to that of Form A, was released in 1971.

Introduction

Every child who enters a school system comes with a different background of experience, and thus starts out with a different body of knowledge and set of understandings. It is essential that educators focus their attention on such variations in cognitive development, with respect to their nature, origins, modifiability, and implications for future achievement.

Unfortunately, concern with these issues is not always reflected in preschool or primary-grade materials. Some authors assume that all beginning pupils possess a fundamental knowledge of certain basic concepts. Others may acknowledge the problem, but provide only a minimum of step-by-step instruction for the child, or else leave it to the teacher to identify and remedy such early differences. The "Directions" portions of instructional materials may use such basic concepts as *below*, *different*, *middle*, *more*, *top*, and *last*, to name but a few, while the accompanying text frequently provides minimal exploration of the meanings of these concepts, and no techniques for assessing children's knowledge of their meanings.

The assumption that children have mastered the basic concepts necessary for understanding and following directions by the time of school entry needs to be questioned. Indeed, studies of reasonably typical samples have shown that as many as 60 out of 100 children entering kindergarten may be unable to mark the *right end* of a line, or to indicate the area *below* a pictured table.[1]

Such findings are justifiably a source of concern because of their possible implications for initial and hence subsequent school achievement. Research results indicate not only that the pupil who starts out behind tends to stay behind, but also that the gap between good and poor achievers tends to become wider and wider over time.[2] Because this suggests that the effects of beginning pupils' deficiencies may be cumulative, attention is being devoted increasingly to the early correction of initial lags in concept and language development—lags that in many cases may be traceable to preschool deprivation in such learning experiences.

It is within this context of recent research and practice that the *Boehm Test of Basic Concepts* (*BTBC*) was conceived, and it is to the joint problems of detection and remediation that it is addressed.

Development and Purpose

The items comprising Form A of the *BTBC* were developed by a multi-stage process. First, a comprehensive selection of preschool and primary-grade curriculum materials in the areas of reading, arithmetic, and science was reviewed. From the directions and other portions of these materials, terms were selected which (a) occurred with considerable frequency; (b) were seldom if ever explicitly defined, or were defined in their simple forms but subsequently used in complex forms without adequate transitions; and (c) represented relatively abstract basic concepts or ideas.

[1] Based on data from pre-standardization tryout of experimental items for the *Boehm Test of Basic Concepts*, Form A.

[2] A. E. Boehm, *The development of comparative concepts in primary school children*. (Doctoral dissertation, Columbia University) Ann Arbor, Michigan: University Microfilms, 1967. No. 67-5767.

Once chosen, the concepts fulfilling these criteria were translated into pictorial, multiple-choice items, and the items were tried out twice on appropriate groups. The purpose of these tryouts (the details of which are described in the Technical Data section of this Manual) was to identify (a) items that children found difficult for reasons other than lack of concept knowledge, such as ambiguous wording or unclear art work, and (b) items that were answered correctly by nearly all (98 per cent or more) of the kindergarten and first-grade pupils tested. Both of these sets of items were eliminated—the former for obvious reasons and the latter on the grounds that teachers naturally wish to focus attention not on concepts already familiar to their classes, but on those which children find difficult. Of course, there are some children who are unfamiliar with these "easy" concepts. To include such concepts in the *BTBC*, however, would have entailed either lengthening the test or omitting items considered more important for the test in the sense that they relate to concepts less familiar, on the average, to preschool and primary-grade children. Thus, another criterion—"concept unfamiliarity"—was utilized to refine further the selection of items for the test.[3]

Items for Form B were written to parallel those included on Form A, with like-numbered items on the two forms measuring knowledge of the same basic concept.

The rules used to select the concepts tested by the *BTBC* define both the rationale and the purpose of the test. The instrument is designed to assess beginning school children's knowledge of frequently used basic concepts widely but sometimes mistakenly assumed to be familiar to children at their time of entry into kindergarten or first grade. As such, the purpose of the *BTBC* is to identify:

(a) individual *children* whose overall level of concept mastery is low and who therefore may need special attention, and

(b) individual *concepts* with which large numbers of children in a class may be unfamiliar.

In this way the test is intended both as a "detector" and as an instructional device for use by the classroom teacher.

[3] It should be noted that "concept unfamiliarity" refers here only to the child's ability to respond to concepts *as presented on the printed page*. It is possible that some children would be unable to answer certain items correctly in the test booklet but could respond correctly when comparing physical objects. However, since children in a school situation are dealing with printed materials, the *BTBC* provides the teacher with relevant information in pointing out weaknesses in this area.

ADMINISTRATION AND SCORING

How to Administer the Test

Administering the *Boehm Test of Basic Concepts* is simple and straightforward. Form A and Form B are administered in the same manner. Each of the forms consists of 50 pictorial items arranged in approximate order of increasing difficulty and divided evenly between two booklets, each containing 3 sample questions followed by 25 test questions. Booklet 2 is more difficult than Booklet 1. Each item consists of a set of pictures, about which statements are read aloud to the children by the examiner. The statements briefly describe the pictures and instruct the children to mark the one that illustrates the concept being tested.

Each of the booklets requires approximately 15 to 20 minutes to administer to kindergarten classes. This time estimate includes the time needed for giving general instructions and reviewing the sample questions. Older groups should require somewhat less time than younger classes.

Experience with the tryout and standardization groups has indicated that most children become interested in the test and enjoy taking it, so that holding the children's attention, even at the kindergarten level, is usually not difficult. Both booklets of a given form may be administered at a single session if the children remain attentive. However, there is considerable advantage in administering the *BTBC* in two separate sessions if the children are very young or lack the attention span required to respond alertly to all 50 items of the test at a single sitting. *This decision should be made before starting the second booklet, since a booklet once begun should be completed in the same session.* A booklet should not be split between two sessions.

If very young children who are unfamiliar with handling and marking workbooks are being tested, small groups of 8 to 12 children are generally advisable. When large groups are being tested, it is helpful to have teachers' aides or proctors assist the children to write their names, check responses to the sample questions, turn pages and find their place, and so forth. (Alternatively, the teacher may wish to write the children's names on the booklets in advance and make a corresponding adjustment in the directions given to the children.)

The testing materials consist of (a) the Manual and a demonstration copy of both test booklets for the teacher, and (b) two test booklets plus a crayon or pencil for each child.

Before beginning the test, seat the children as far apart as possible, and give each pupil a pencil or crayon. Using the demonstration booklet, begin with the 3 sample questions. When these have been completed, check each child's responses and have him correct any errors before proceeding to the test questions.

Detailed directions are provided below for each booklet of Form A and each booklet of Form B. These should be followed exactly, with the **boldface** portions being read aloud to the children. The key phrases in each item are read *twice*, with emphasis on the ***italicized*** words.

Allow the children enough time to respond to each question before going on to the next. With very young children, it may be necessary to help them keep their place by indicating the appropriate set of pictures in the demonstration booklet as each question is read. Be careful, however, not to point to any particular picture within a set.

The children may correct errors on either the sample or the test questions by erasing or by encircling the incorrect answer, and marking the new answer in the regular way.

Spanish translations of the directions for administration for Forms A and B of the *BTBC* are available in separate pamphlets. Order Catalog No. **51,466** for Form A, or Catalog No. **51,480** for Form B.

Form A: Detailed Directions

Booklet 1

When ready to begin testing, say to the children:

"I am going to give each of you a book. Leave it on your desk until I tell you what to do."

Distribute copies of Booklet 1, *face up*. Then say:

"I have given you a book with some pictures in it. We are going to do different kinds of things with the pictures. Listen and do just what I say. First, print your name on the line up here."

Point to the line at the top of the cover of the demonstration copy. (With very young children, it may be necessary to write the names for the children.)

Sample Questions

When the names have been written, say:

"We are going to look at pictures and mark X's on them. This is how you make an X."

Draw a large X on the blackboard. Then say:

"Now find the gray box with the telephone in it. Put your finger on it."

Check to see that every child has found the gray box. Assist those who are having trouble. When each pupil has his finger on the box, start reading the sample questions. Remember to emphasize the italicized words. Begin by saying:

"Now take your finger off the gray box and pick up your pencil (or crayon).

"Now look at the shoe, the hat, and the sock. Mark an X on the *hat*. . . . Mark an X right on the *hat*."

Wait until all of the children have responded. Then say:

"Now look at the things to ride in. Mark an X on the *boat*. . . . Mark the *boat*.

"Look at the fruit. Mark the *banana*. . . . Mark the *banana*.

"Very good. Now put down your pencils (or crayons). I will look at your books. Do *not* turn the page. If you make a mistake or want to change an answer, make a circle around it like this (demonstrate on the blackboard) **and then make the new mark."**

Make certain that each child has written his name correctly and has put X's on the hat, the boat, and the banana. Correct the child's name where necessary. If anyone has marked a wrong item, point out the error and have the child correct it. If any child's X's are not *directly* on the hat, the boat, or the banana, ask him to make these corrections also.

When all work has been checked, start to read the Test Questions.

Test Questions

Say:

"Now open your books."

Assist the children if necessary. Then point to the gray box on the left-hand page of the demonstration booklet and say:

"Look for the gray box like this one on your page. Put your finger on the gray box with the book in it."

Check to see that each child has found the proper box. Then start reading the test questions. (Do *not* read the question numbers.)

1. **"Now take your finger off the box and pick up your pencil (or crayon). Look at the pictures of writing paper with stars. Mark the paper with the star at the *top*. . . . Mark the paper with the star at the *top*.**

2. **"Look at the beads and strings. Mark the bead that has a string *through* it. . . . Mark the bead that has a string *through* it.**

3. **"Look at the table and the boxes. Mark the box that is *away from* the table. . . . Mark the box that is *away from* the table.**

4. **"Look at the toys. Mark the toy that is *next to* the truck. . . . Mark the toy that is *next to* the truck."**

Then point to the gray box on the right-hand page of your demonstration booklet, and say:

"Now put your finger on the gray box with the candle in it."

See that everyone has found the proper box. Then say:

"Now pick up your pencil.

5. **"Look at the pictures of the house and the boy. Mark the house with the boy *inside* it. . . . Mark the house with the boy *inside* it.**

6. **"Look at the boxes and marbles. Mark the box that has *some but not many* marbles. . . . Mark the box that has *some but not many* marbles.**

7. "Look at the flowers. Mark the flower that is in the *middle*. . . . Mark the flower that is in the *middle*.

8. "Look at the plates of cupcakes. Mark the plate that has a *few* cupcakes. . . . Mark the plate that has a *few* cupcakes.

"Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

9. "Look at the boats. Mark the boat that is *farthest* from the shore. . . . Mark the boat that is *farthest* from the shore.

10. "Look at the boxes and circles. Mark the box that has circles *around* it. . . . Mark the box that has circles *around* it.

11. "Look at the balloons and the tree. Mark the balloon that is *over* the tree. . . . Mark the balloon that is *over* the tree.

"Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

12. "Look at the doors. Mark the door that is *widest*. . . . Mark the door that is *widest*.

13. "Look at the boxes of eggs. Mark the box that has the *most* eggs. . . . Mark the box that has the *most* eggs.

14. "Look at the jars, cups, and spoons. Mark the thing that is *between* the spoons. . . . Mark the thing that is *between* the spoons.

15. "Look at the cakes. Mark the cake that is *whole*. . . . Mark the cake that is *whole*.

"Now turn the page. Find the gray box with the hat in it."

Demonstrate and point to the box.

16. "Look at the boys going to school. Mark the boy who is *nearest* the door. . . . Mark the boy who is *nearest* the door.

17. "Look at the animals walking in a line. Mark the *second* animal. . . . Mark the *second* animal.

18. "Look at the glasses on the table. Mark the glass that is at a *corner of* the table. . . . Mark the glass that is at a *corner of* the table.

"Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

19. "Look at the groups of animals. Mark the group that has *several* rabbits. . . . Mark the group that has *several* rabbits.

20. "Look at the sofa and the toys. Mark the toy that is *behind* the sofa. . . . Mark the toy that is *behind* the sofa.

21. "Look at the groups of trees. Mark the group where all the trees are in a *row*. . . . Mark the group where all the trees are in a *row*.

"Now turn the page. Find the gray box with the bottle in it."

Demonstrate and point to the box.

22. "Look at the groups of blocks. Mark the group that is *different* from the others. . . . Mark the group that is *different* from the others.

23. "Look at the pictures of a girl. Mark the picture that shows how the girl looked *after* her hair was cut. . . . Mark the picture that shows how the girl looked *after* her hair was cut.

24. "Look at the bottles. Mark the one that is *almost* empty. . . . Mark the one that is *almost* empty.

25. "Look at the pies. Mark the pie that is *half* gone. . . . Mark the pie that is *half* gone."

Form A: Detailed Directions

Booklet 2

When ready to begin testing, say to the children:

"I am going to give each of you another book. Leave it on your desk until I tell you what to do."

Distribute copies of Booklet 2, *face up*. Then say:

"I have given you another book of pictures. Listen and do just what I say. First, print your name on the line up here."

Point to the line at the top of the cover of the demonstration copy. Write the children's names for them, if necessary.

Sample Questions

When the names have been written, say:

"We are going to mark X's on pictures as we did before. Remember, this is how you make an X."

Draw a large X on the blackboard. Then say:

"Now find the gray box with the telephone in it. Put your finger on it."

Check to see that every child has found the gray box. When everyone has his finger on the box, begin reading the sample questions. Start by saying:

"Now take your finger off the gray box and pick up your pencil (or crayon).

"Now look at the spoon, glass, and cup. Mark an X on the *glass*. . . . Mark an X on the *glass*."

When all of the children have responded, say:

"Look at the furniture. Mark an X on the *table*. . . . Mark an X on the *table*.

"Now look at the animals. Mark the *dog*. . . . Mark the *dog*.

"Very good. Now put down your pencils (or crayons). I will look at your books. Do *not* turn the page."

Make certain that each child has written his name correctly and put X's directly on the glass, the table, and the dog. Have them make any necessary corrections.

When all of the booklets have been checked, start to read the Test Questions.

Test Questions

Say:

"Now open your books."

Assist the children if necessary. Then point to the gray box on the left-hand page of your demonstration booklet and say:

"Look for the gray box like this one on your page. Put your finger on the gray box with the book in it."

Check to see that each child has found the proper box. Then read the test questions. (Do *not* read the question numbers.)

26. "Now take your finger off the box and pick up your pencil (or crayon). Look at the circle and the boxes. Mark the box that is at the *center* of the circle. . . . Mark the box that is at the *center* of the circle.

27. "Look at the box of marbles and the groups of marbles. Mark the group that has *as many* marbles as the box. . . . Mark the group that has *as many* marbles as the box.

28. "Look at the box and the circles. Mark the circle that is at a *side* of the box. . . . Mark the circle that is at a *side* of the box."

Point to the gray box on the right-hand page of the demonstration booklet, and say:

"Now find the gray box with the candle in it.

29. "Look at the trees and squirrels. Mark the squirrel that is *beginning* to climb a tree. . . . Mark the squirrel that is *beginning* to climb a tree.

30. "Look at the desserts. One is an ice cream cone and one is a piece of pie. Mark the *other* dessert. . . . Mark the *other* dessert.

31. "Look at the shapes. Mark the shapes that are *alike*. . . . Mark the shapes that are *alike*.

32. "Look at the cars going into the tunnel. Mark the car that is *not the first or the last*. . . . Mark the car that is *not the first or the last*.

"Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

33. "Look at the chair, the apple, and the cookies. Mark what a child should *never* eat. . . . Mark what a child should *never* eat.

34. "Look at the table. Make an X *below* the table. . . . Make an X *below* the table.

35. "Look at the boxes and the balls. Mark the ball that *matches* one of the boxes. . . . Mark the ball that *matches* one of the boxes.

36. "Look at the dog, the book, and the ear. Mark the one a child *always* has. . . . Mark the one a child *always* has.

"Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

37. "Look at the fish. Mark the fish that is *medium-sized*. . . . Mark the fish that is *medium-sized*.

38. "Look at the boxes and the line. Mark the box that is over the *right* end of the line. . . . Mark the box that is over the *right* end of the line.

39. "Look at the boys. Mark the boy who is bending *forward*. . . . Mark the boy who is bending *forward*.

40. "Look at the boxes and candies. Mark the box that has *zero* candies. . . . Mark the box that has *zero* candies.

"Now turn the page. Find the gray box with the hat in it."

Demonstrate.

41. "Look at the cloud and the airplanes. Mark the airplane that is *above* the cloud. . . . Mark the airplane that is *above* the cloud.

42. "Look at the pictures of bowls and spoons. Mark the picture that shows a spoon in *every* bowl. . . . Mark the picture that shows a spoon in *every* bowl.

43. "Look at the beads. Mark the beads that are *separated*. . . . Mark the beads that are *separated*.

"Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

44. "Look at the birds. Mark the bird on the *left*. . . . Mark the bird on the *left*.

45. "Look at the pictures of candles. Mark the picture that shows a *pair* of candles. . . . Mark the picture that shows a *pair* of candles.

46. "Look at the boxes. One box has an X in it. *Skip* a box and make another X. . . . *Skip* a box and make another X.

47. "Look at the pictures of lollipops. Mark the pictures that have *equal* numbers of lollipops. . . . Mark the pictures that have *equal* numbers of lollipops.

"Now turn the page. Find the gray box with the bottle in it."

Demonstrate.

48. "Look at the boxes of circles. Mark the box where the circles are *in order* from large to small. . . . Mark the box where the circles are *in order* from large to small.

49. "Look at the teacher and the children. Mark the *third* child from the teacher. . . . Mark the *third* child from the teacher.

50. "Look at the groups of stars. Mark the group that has the *least* stars. . . . Mark the group that has the *least* stars."

7. **"Look at the flowers. Mark the flower that is in the *middle*. . . . Mark the flower that is in the *middle*.**

8. **"Look at the plates of cupcakes. Mark the plate that has a *few* cupcakes. . . . Mark the plate that has a *few* cupcakes.**

"Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

9. **"Look at the boats. Mark the boat that is *farthest* from the shore. . . . Mark the boat that is *farthest* from the shore.**

10. **"Look at the boxes and circles. Mark the box that has circles *around* it. . . . Mark the box that has circles *around* it.**

11. **"Look at the balloons and the tree. Mark the balloon that is *over* the tree. . . . Mark the balloon that is *over* the tree.**

"Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

12. **"Look at the doors. Mark the door that is *widest*. . . . Mark the door that is *widest*.**

13. **"Look at the boxes of eggs. Mark the box that has the *most* eggs. . . . Mark the box that has the *most* eggs.**

14. **"Look at the jars, cups, and spoons. Mark the thing that is *between* the spoons. . . . Mark the thing that is *between* the spoons.**

15. **"Look at the cakes. Mark the cake that is *whole*. . . . Mark the cake that is *whole*.**

"Now turn the page. Find the gray box with the hat in it."

Demonstrate and point to the box.

16. **"Look at the boys going to school. Mark the boy who is *nearest* the door. . . . Mark the boy who is *nearest* the door.**

17. **"Look at the animals walking in a line. Mark the *second* animal. . . . Mark the *second* animal.**

18. **"Look at the glasses on the table. Mark the glass that is at a *corner* of the table. . . . Mark the glass that is at a *corner* of the table.**

"Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

19. **"Look at the groups of animals. Mark the group that has *several* rabbits. . . . Mark the group that has *several* rabbits.**

20. **"Look at the sofa and the toys. Mark the toy that is *behind* the sofa. . . . Mark the toy that is *behind* the sofa.**

21. **"Look at the groups of trees. Mark the group where all the trees are in a *row*. . . . Mark the group where all the trees are in a *row*.**

"Now turn the page. Find the gray box with the bottle in it."

Demonstrate and point to the box.

22. **"Look at the groups of blocks. Mark the group that is *different* from the others. . . . Mark the group that is *different* from the others.**

23. **"Look at the pictures of a girl. Mark the picture that shows how the girl looked *after* her hair was cut. . . . Mark the picture that shows how the girl looked *after* her hair was cut.**

24. **"Look at the bottles. Mark the one that is *almost* empty. . . . Mark the one that is *almost* empty.**

25. **"Look at the pies. Mark the pie that is *half* gone. . . . Mark the pie that is *half* gone."**

Form A: Detailed Directions

Booklet 2

When ready to begin testing, say to the children:

"I am going to give each of you another book. Leave it on your desk until I tell you what to do."

Distribute copies of Booklet 2, *face up*. Then say:

"I have given you another book of pictures. Listen and do just what I say. First, print your name on the line up here."

Point to the line at the top of the cover of the demonstration copy. Write the children's names for them, if necessary.

Sample Questions

When the names have been written, say:

"We are going to mark X's on pictures as we did before. Remember, this is how you make an X."

Draw a large X on the blackboard. Then say:

"Now find the gray box with the telephone in it. Put your finger on it."

Check to see that every child has found the gray box. When everyone has his finger on the box, begin reading the sample questions. Start by saying:

"Now take your finger off the gray box and pick up your pencil (or crayon).

"Now look at the spoon, glass, and cup. Mark an X on the *glass*. . . . Mark an X on the *glass*."

When all of the children have responded, say:

"Look at the furniture. Mark an X on the *table*. . . . Mark an X on the *table*.

"Now look at the animals. Mark the *dog*. . . . Mark the *dog*.

"Very good. Now put down your pencils (or crayons). I will look at your books. Do *not* turn the page."

Make certain that each child has written his name correctly and put X's directly on the glass, the table, and the dog. Have them make any necessary corrections.

When all of the booklets have been checked, start to read the Test Questions.

Test Questions

Say:

"Now open your books."

Assist the children if necessary. Then point to the gray box on the left-hand page of your demonstration booklet and say:

"Look for the gray box like this one on your page. Put your finger on the gray box with the book in it."

Check to see that each child has found the proper box. Then read the test questions. (Do *not* read the question numbers.)

26. **"Now take your finger off the box and pick up your pencil (or crayon). Look at the circle and the boxes. Mark the box that is at the *center* of the circle. . . . Mark the box that is at the *center* of the circle.**

27. **"Look at the box of marbles and the groups of marbles. Mark the group that has *as many* marbles as the box. . . . Mark the group that has *as many* marbles as the box.**

28. **"Look at the box and the circles. Mark the circle that is at a *side* of the box. . . . Mark the circle that is at a *side* of the box."**

Point to the gray box on the right-hand page of the demonstration booklet, and say:

"Now find the gray box with the candle in it.

29. **"Look at the trees and squirrels. Mark the squirrel that is *beginning* to climb a tree. . . . Mark the squirrel that is *beginning* to climb a tree.**

30. **"Look at the desserts. One is an ice cream cone and one is a piece of pie. Mark the *other* dessert. . . . Mark the *other* dessert.**

31. **"Look at the shapes. Mark the shapes that are *alike*. . . . Mark the shapes that are *alike*.**

32. **"Look at the cars going into the tunnel. Mark the car that is *not the first or the last*. . . . Mark the car that is *not the first or the last*.**

"Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

33. **"Look at the chair, the apple, and the cookies. Mark what a child should *never* eat. . . . Mark what a child should *never* eat.**

34. **"Look at the table. Make an X *below* the table. . . . Make an X *below* the table.**

35. **"Look at the boxes and the balls. Mark the ball that *matches* one of the boxes. . . . Mark the ball that *matches* one of the boxes.**

36. **"Look at the dog, the book, and the ear. Mark the one a child *always* has. . . . Mark the one a child *always* has.**

"Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

37. **"Look at the fish. Mark the fish that is *medium-sized*. . . . Mark the fish that is *medium-sized*.**

38. **"Look at the boxes and the line. Mark the box that is over the *right* end of the line. . . . Mark the box that is over the *right* end of the line.**

39. **"Look at the boys. Mark the boy who is bending *forward*. . . . Mark the boy who is bending *forward*.**

40. **"Look at the boxes and candies. Mark the box that has *zero* candies. . . . Mark the box that has *zero* candies.**

"Now turn the page. Find the gray box with the hat in it."

Demonstrate.

41. **"Look at the cloud and the airplanes. Mark the airplane that is *above* the cloud. . . . Mark the airplane that is *above* the cloud.**

42. **"Look at the pictures of bowls and spoons. Mark the picture that shows a spoon in *every* bowl. . . . Mark the picture that shows a spoon in *every* bowl.**

43. **"Look at the beads. Mark the beads that are *separated*. . . . Mark the beads that are *separated*.**

"Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

44. **"Look at the birds. Mark the bird on the *left*. . . . Mark the bird on the *left*.**

45. **"Look at the pictures of candles. Mark the picture that shows a *pair* of candles. . . . Mark the picture that shows a *pair* of candles.**

46. **"Look at the boxes. One box has an X in it. *Skip* a box and make another X. . . . *Skip* a box and make another X.**

47. **"Look at the pictures of lollipops. Mark the pictures that have *equal* numbers of lollipops. . . . Mark the pictures that have *equal* numbers of lollipops.**

"Now turn the page. Find the gray box with the bottle in it."

Demonstrate.

48. **"Look at the boxes of circles. Mark the box where the circles are *in order* from large to small. . . . Mark the box where the circles are *in order* from large to small.**

49. **"Look at the teacher and the children. Mark the *third* child from the teacher. . . . Mark the *third* child from the teacher.**

50. **"Look at the groups of stars. Mark the group that has the *least* stars. . . . Mark the group that has the *least* stars."**

Form B: Detailed Directions

Booklet 1

When ready to begin testing, say to the children:

"I am going to give each of you a book. Leave it on your desk until I tell you what to do."

Distribute copies of Booklet 1, *face up*. Then say:

"I have given you a book with some pictures in it. We are going to do different kinds of things with the pictures. Listen and do just what I say. First, print your name on the line up here."

Point to the line at the top of the cover of the demonstration copy. (With very young children, it may be necessary to write the names for the children.)

Sample Questions

When the names have been written, say:

"We are going to look at pictures and mark X's on them. This is how you make an X."

Draw a large X on the blackboard. Then say:

"Now find the gray box with the telephone in it. Put your finger on it."

Check to see that every child has found the gray box. Assist those who are having trouble. When each pupil has his finger on the box, start reading the sample questions. Remember to emphasize the italicized words. Begin by saying:

"Now take your finger off the gray box and pick up your pencil (or crayon).

"Now look at the shoe, the hat, and the sock. Mark an X on the *hat*. . . . Mark an X right on the *hat*."

Wait until all of the children have responded. Then say:

"Now look at the things to ride in. Mark an X on the *boat*. . . . Mark the *boat*.

"Look at the fruit. Mark the *banana*. . . . Mark the *banana*.

"Very good. Now put down your pencils (or crayons). I will look at your books. Do *not* turn the page. If you make a mistake or want to change an answer, make a circle around it like this (demonstrate on the blackboard) **and then make the new mark."**

Make certain that each child has written his name correctly and has put X's on the hat, the boat, and the banana. Correct the child's name where necessary. If anyone has marked a wrong item, point out the error and have the child correct it. If any child's X's are not *directly* on the hat, the boat, or the banana, ask him to make these corrections also.

When all work has been checked, start to read the Test Questions.

Test Questions

Say:

"Now open your books."

Assist the children if necessary. Then point to the gray box on the left-hand page of the demonstration booklet and say:

"Look for the gray box like this one on your page. Put your finger on the gray box with the book in it."

Check to see that each child has found the proper box. Then start reading the test questions. (Do *not* read the question numbers.)

1. **"Now take your finger off the box and pick up your pencil (or crayon). Look at the flags on the poles. Mark the pole with the flag at the *top*. . . . Mark the pole with the flag at the *top*.**
2. **"Look at the dogs and the hoops. Mark the dog that is going *through* the hoop. . . . Mark the dog that is going *through* the hoop.**
3. **"Look at the baby and the blocks. Mark the block that is *away from* the baby. . . . Mark the block that is *away from* the baby."**

Then point to the gray box on the right-hand page of your demonstration booklet, and say:

"Now put your finger on the gray box with the candle in it."

See that everyone has found the proper box. Then say:

"Now pick up your pencil.

4. **"Look at the animals. Mark the animal that is *next to* the rabbit. . . . Mark the animal that is *next to* the rabbit.**
5. **"Look at the boxes and balls. Mark the box with the balls *inside* it. . . . Mark the box with the balls *inside* it.**
6. **"Look at the bowls of flowers. Mark the bowl that has *some but not many* flowers. . . . Mark the bowl that has *some but not many* flowers.**
7. **"Look at the children. Mark the child who is in the *middle*. . . . Mark the child who is in the *middle*.**

   "Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

8. **"Look at the pictures of boxes. Mark the picture that has a *few* boxes. . . . Mark the picture that has a *few* boxes.**
9. **"Look at the clothes hanging on the line. Mark the dress that is *farthest* from the socks. . . . Mark the dress that is *farthest* from the socks.**
10. **"Look at the flowers and strings. Mark the flower that has a string *around* it. . . . Mark the flower that has a string *around* it.**

11. **"Look at the children and the rope. Mark the child who is *over* the rope. . . . Mark the child who is *over* the rope.**

    "Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

12. **"Look at the ties. Mark the tie that is *widest*. . . . Mark the tie that is *widest*.**
13. **"Look at the boxes of buttons. Mark the box that has the *most* buttons. . . . Mark the box that has the *most* buttons.**
14. **"Look at the pictures of toys. Mark the picture that has a bear *between* two blocks. . . . Mark the picture that has a bear *between* two blocks.**
15. **"Look at the apples. Mark the apple that is *whole*. . . . Mark the apple that is *whole*.**

    "Now turn the page. Find the gray box with the hat in it."

Demonstrate and point to the box.

16. **"Look at the dogs and the bone. Mark the dog that is *nearest* the bone. . . . Mark the dog that is *nearest* the bone.**
17. **"Look at the line of trucks and the sign. Mark the *second* truck from the sign. . . . Mark the *second* truck from the sign.**
18. **"Look at the buildings. Mark the building that is at a *corner* of the street. . . . Mark the building that is at a *corner* of the street.**

    "Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

19. **"Look at the groups of knives, forks, and spoons. Mark the group that has *several* spoons. . . . Mark the group that has *several* spoons.**
20. **"Look at the boys and the wagon. Mark the boy who is *behind* the wagon. . . . Mark the boy who is *behind* the wagon.**
21. **"Look at the pictures of bottles. Mark the picture where all the bottles are in a *row*. . . . Mark the picture where all the bottles are in a *row*.**

    "Now turn the page. Find the gray box with the bottle in it."

Demonstrate and point to the box.

22. **"Look at the piles of books. Mark the pile that is *different* from the others. . . . Mark the pile that is *different* from the others.**
23. **"Look at the pictures of a piece of wood. Mark the picture that shows how the wood looked *after* it was cut. . . . Mark the picture that shows how the wood looked *after* it was cut.**
24. **"Look at the baskets of fruit. Mark the basket that is *almost* full. . . . Mark the basket that is *almost* full.**
25. **"Look at the boxes. Mark the box that is *half* black. . . . Mark the box that is *half* black."**

Form B: Detailed Directions

Booklet 2

When ready to begin testing, say to the children:

"I am going to give each of you another book. Leave it on your desk until I tell you what to do."

Distribute copies of Booklet 2, *face up*. Then say:

"I have given you another book of pictures. Listen and do just what I say. First, print your name on the line up here."

Point to the line at the top of the cover of the demonstration copy. Write the children's names for them, if necessary.

Sample Questions

When the names have been written, say:

"We are going to mark X's on pictures as we did before. Remember, this is how you make an X."

Draw a large X on the blackboard. Then say:

"Now find the gray box with the telephone in it. Put your finger on it."

Check to see that every child has found the gray box. When everyone has his finger on the box, begin reading the sample questions. Start by saying:

"Now take your finger off the gray box and pick up your pencil (or crayon).

"Now look at the spoon, glass, and cup. Mark an X on the *glass*. . . . Mark an X on the *glass*."

When all of the children have responded, say:

"Look at the furniture. Mark an X on the *table*. . . . Mark an X on the *table*.

"Now look at the animals. Mark the *dog*. . . . Mark the *dog*.

"Very good. Now put down your pencils (or crayons). I will look at your books. Do *not* turn the page."

Make certain that each child has written his name correctly and put X's directly on the glass, the table, and the dog. Have them make any necessary corrections.

When all of the booklets have been checked, start to read the Test Questions.

Test Questions

Say:

"Now open your books."

Assist the children if necessary. Then point to the gray box on the left-hand page of your demonstration booklet and say:

"Look for the gray box like this one on your page. Put your finger on the gray box with the book in it."

Check to see that each child has found the proper box. Then read the test questions. (Do *not* read the question numbers.)

26. "Now take your finger off the box and pick up your pencil (or crayon). Look at the ring and the marbles. Mark the marble that is at the *center* of the ring. . . . Mark the marble that is at the *center* of the ring.

27. "Look at the box of pencils and the groups of pencils. Mark the group that has *as many* pencils as the box. . . . Mark the group that has *as many* pencils as the box.

28. "Look at the car and the boys. Mark the boy at the *side* of the car. . . . Mark the boy at the *side* of the car."

Point to the gray box on the right-hand page of the demonstration booklet, and say:

"Now find the gray box with the candle in it.

29. "Look at the boys on the stairs. Mark the boy who is *beginning* to climb the stairs. . . . Mark the boy who is *beginning* to climb the stairs.

30. "Look at the toys. One is a doll and one is a truck. Mark the *other* toy. . . . Mark the *other* toy.

31. "Look at the socks. Mark the socks that are *alike*. . . . Mark the socks that are *alike*.

32. "Look at the ducks in the water. Mark the duck that is *not the first or the last*. . . . Mark the duck that is *not the first or the last*.

"Now turn the page and find the gray box with the scissors in it."

Demonstrate and point to the box.

33. "Look at the lamp, the wristwatch, and the shoe. Mark the thing that a child should *never* wear. . . . Mark the thing that a child should *never* wear.

34. "Look at the bench and the birds. Mark the bird that is *below* the bench. . . . Mark the bird that is *below* the bench.

35. "Look at the shirts and pants. Mark the pants that *match* one of the shirts. . . . Mark the pants that *match* one of the shirts.

"Now find the gray box with the pencil in it."

Point to the box on the right-hand page.

36. "Look at the box, the wheel, and the feather. Mark the thing a bicycle *always* has. . . . Mark the thing a bicycle *always* has.

37. "Look at the butterflies. Mark the butterfly that is *medium-sized*. . . . Mark the butterfly that is *medium-sized*.

38. "Look at the apples on the shelf. Mark the apple at the *right* end of the shelf. . . . Mark the apple at the *right* end of the shelf.

39. "Look at the little chicks. Mark the chick that is bending *forward*. . . . Mark the chick that is bending *forward*.

"Now turn the page. Find the gray box with the hat in it."

Demonstrate.

40. "Look at the rabbits and carrots. Mark the rabbit that has *zero* carrots. . . . Mark the rabbit that has *zero* carrots.

41. "Look at the windows of the house. Mark the window that is *above* the door. . . . Mark the window that is *above* the door.

42. "Look at the groups of circles and dots. Mark the group that has a dot in *every* circle. . . . Mark the group that has a dot in *every* circle.

"Now find the gray box with the light bulb in it."

Point to the box on the right-hand page.

43. "Look at the pictures of boxes. Mark the picture where the boxes are *separated*. . . . Mark the picture where the boxes are *separated*.

44. "Look at the trees. Mark the tree on the *left*. . . . Mark the tree on the *left*.

45. "Look at the pictures of dolls. Mark the picture that shows a *pair* of dolls. . . . Mark the picture that shows a *pair* of dolls.

46. "Look at the circles. One circle has an X in it. *Skip* a circle and make another X. . . . *Skip* a circle and make another X.

"Now turn the page. Find the gray box with the bottle in it."

Demonstrate.

47. "Look at the groups of stars. Mark the groups that have *equal* numbers of stars. . . . Mark the groups that have *equal* numbers of stars.

48. "Look at the pictures of boxes. Mark the picture where the boxes are *in order* from small to large. . . . Mark the picture where the boxes are *in order* from small to large.

49. "Look at the store and the houses. Mark the *third* house from the store. . . . Mark the *third* house from the store.

50. "Look at the pictures of ice cream cones. Mark the picture that has the *least* cones. . . . Mark the picture that has the *least* cones."

How to Score and Record the Results

The Class Record Form serves as both a scoring key and an interpretive aid to the teacher. It is designed to speed recording and facilitate analysis of a class's test results. If there are more than 30 children in the class, two Class Record Forms should be fastened together so that the entries in the shaded column headed "Total Number of Children Answering Correctly" and the adjacent column headed "Per Cent Passing" will be based on the scores of the entire class.

Down the left-hand side of the form the concepts covered by the test are listed by item number, in the sequence in which they appear in the test booklets. Each concept is accompanied by a miniature reproduction of the set of pictures by which it is tested, appropriately marked to show the correct response. The Class Record Forms for both forms of the test are alike with respect to the concepts listed and the sequence of listing but of course they have different picture reproductions, each marked with the correct response. An illustration of the Form A Class Record Form is provided on page 13.

Across the top of the form are spaces for entering the names of the children tested. The form has been marked off into boxes for recording the response of each child to each item. If the same children are given both forms of the test (for example, Form A at the beginning and Form B at the end of the school year), it should prove helpful in comparing the results of the two testings if the names of the children have been entered in the same sequence on both Class Record Forms.

Two suggested procedures for using the Class Record Form are described below. With either plan it will require approximately one hour to complete the form for a group of 30 children. About 20 additional minutes are needed to cross-check the column and row sums and to enter the per cent passing for each of the 50 items.

Plan A

The following procedure is recommended for scoring Booklets 1 and 2 together:

1. Place together the two booklets for each child (with Booklet 2 following Booklet 1). Then stack the pairs of booklets in any order desired (for example, alphabetically).

2. Write the children's names, in the order in which the booklets are stacked, in the spaces provided for this purpose at the top of the Class Record Form.

3. Starting with the pair of booklets for the first child, go through them item by item, making a check mark (✓) in the appropriate space of that child's column for each item answered *correctly*. Use the miniature scoring key on the Class Record Form for this purpose.[4] Make no check marks at all for items answered incorrectly or ambiguously, or for items that have been omitted. Repeat this procedure for each child tested.

4. Count the check marks in each *row* and record the total in the *shaded column* marked "Total Number of Children Answering Correctly" on the right-hand side of the Class Record Form.

5. Count the check marks in each *column* and record the total in the *shaded row* labeled "Score (Total Number of Items Answered Correctly)" at the bottom of the form.

6. Add the numbers in the *shaded row*—"Score (Total Number of Items Answered Correctly)"—and record this figure in the box labeled "Sum." It should agree with the total of the *shaded column* ("Total Number of Children Answering Correctly").[5]

7. Divide the "Sum" by the total number of children tested, and record the result in the box labeled "Class Average." This quotient will represent the group's average score on the total test.

8. To determine the per cent of children in the class who answered each item correctly, divide each entry in the shaded column labeled "Total Number of Children Answering Correctly" by the total number of children tested, and multiply the result by 100. Record these per cents in the column at the right headed "Per Cent Passing." For classes of 11-40 pupils, the table in the Appendix may be used to determine conveniently the per cent of children in the class who answered an item correctly. Directions are printed on the table.

9. A row for entering percentiles is provided at the bottom of the Class Record Form. These may be obtained from the Norms section of this Manual, or from norms based on local data.

Plan B

If the teacher wishes to score Booklets 1 and 2 *separately*, the following is suggested as an alternative procedure:

1. Arrange all copies of Booklet 1 in the order desired. Then arrange all copies of Booklet 2 in this same order.

2. Write the children's names, in the order in which the Booklet 1's are stacked, in the spaces provided for this purpose at the top of the Class Record Form.

3. Fold the Class Record Form back along the line between items 25 and 26, leaving only the upper part, containing items 1-25, exposed. For each child, enter check marks for correct answers to items 1-25 (Booklet 1) in the manner previously described.

4. To score Booklet 2, fold the Class Record Form so that the fold already made along the line between items 25 and 26 is placed along the line immediately above item 1, leaving only the part containing items 26-50 exposed. Enter check marks for correct answers to items 26-50 (Booklet 2). *Be careful to use the correct column for each child.* If desired, the number appearing at the bottom of each name space on the Class Record Form can be entered on the cover of the two booklets marked by that child.

5. Complete the Class Record Form by following steps 4 through 9 of Plan A.

When the Class Record Form has been completed, the examiner may wish to fold the area containing the last two rows ("Score" and "Percentile") under, and then align this fold with the bottom of the "Child's Name" section at the top of the record form. This will facilitate the reading of each child's score.

A Note on Scoring and Recording Accuracy

Several steps are involved in scoring the *BTBC* and recording the results of testing on the Class Record

[4] Alternatively, the teacher may mark a set of the *BTBC* booklets with the correct answers, and use them in scoring the test.

[5] If not the same, check for counting errors.

Form. Because inaccuracies may occur at any step, a study was conducted of errors made by teachers when scoring and recording their classes' tests. Each of the school districts participating in the standardization of Form A was asked to submit the test booklets of one complete classroom group for each of the three grades tested. (The standardization sample is described in the Technical Data section of this Manual.) Approximately 500 sets of booklets were received, and 200 of these were randomly selected for the study. Of these 200 protocols, 60 included errors of incorrectly scoring a *single item* or of recording a correct response in the *wrong cell* of the Class Record Form—that is, under the wrong child's name or across from an inappropriate item. In 23 of the 60 cases where errors such as these were made, the children's total raw scores were correct (although item difficulty values were affected). Of the remaining 37 cases with errors, only five involved errors of more than two points in a child's total raw score. Since these errors of incorrect scoring or recording were found to affect total scores to such a limited extent, they might be considered minor. However, in view of the nature of this test and its suggested uses, any errors take on a special importance. Therefore, care in scoring and recording the results of the *BTBC* is necessary in order to prevent such errors.

A rare but potentially serious error is that caused by incorrectly summing the item tallies on the Class Record Form to obtain total scores.[6] Errors of this kind can be detected if the cross-check described in step 6 of scoring Plan A is carried out. Because occasional errors in addition may be substantial, this cross-checking should *always* be done.

[6] Only 2.5 per cent of the more than 9700 cases included in the beginning-of-year, Form A standardization sample had total scores incorrectly summed from the tally marks. Errors exceeding two points in either direction of the correct sum occurred in only 0.2 per cent of the cases. The range of errors, however, ran from −20 to +10 of the correct total score.

INTERPRETATION

How to Interpret and Use the Results

How a teacher plans to use the results obtained from the *BTBC* will partly determine the manner in which those results are interpreted. Because the questions of use and interpretation are so closely interrelated, it is necessary to discuss them together.

Individual Remediation, Group Instruction, or Both?

The recommended first step in interpretation is to examine the total scores in the shaded row at the bottom of the Class Record Form, circling those scores suggesting that the pupils in question have an undesirably low level of concept mastery and thus need special help. In deciding whether a given total score is to be regarded as "undesirably low," however, one should consider not only the absolute level of the child's performance (that is, the percentage of the test items he answered correctly), but also his performance relative to that of the rest of his class, and to that of other children of his grade, region, and socioeconomic level.

Thus, if a child answers very few of the 50 items correctly, he would probably benefit from special instruction in the concepts missed regardless of how well or poorly other children performed on the test—since the concepts measured are basic to understanding and following directions in the curricula. Whether or not this child should be singled out for *individual* attention, however, depends upon the overall performance of his class. If the class average on the test is low, the class as a whole is likely to benefit from remedial instruction, and various group training procedures should be considered.

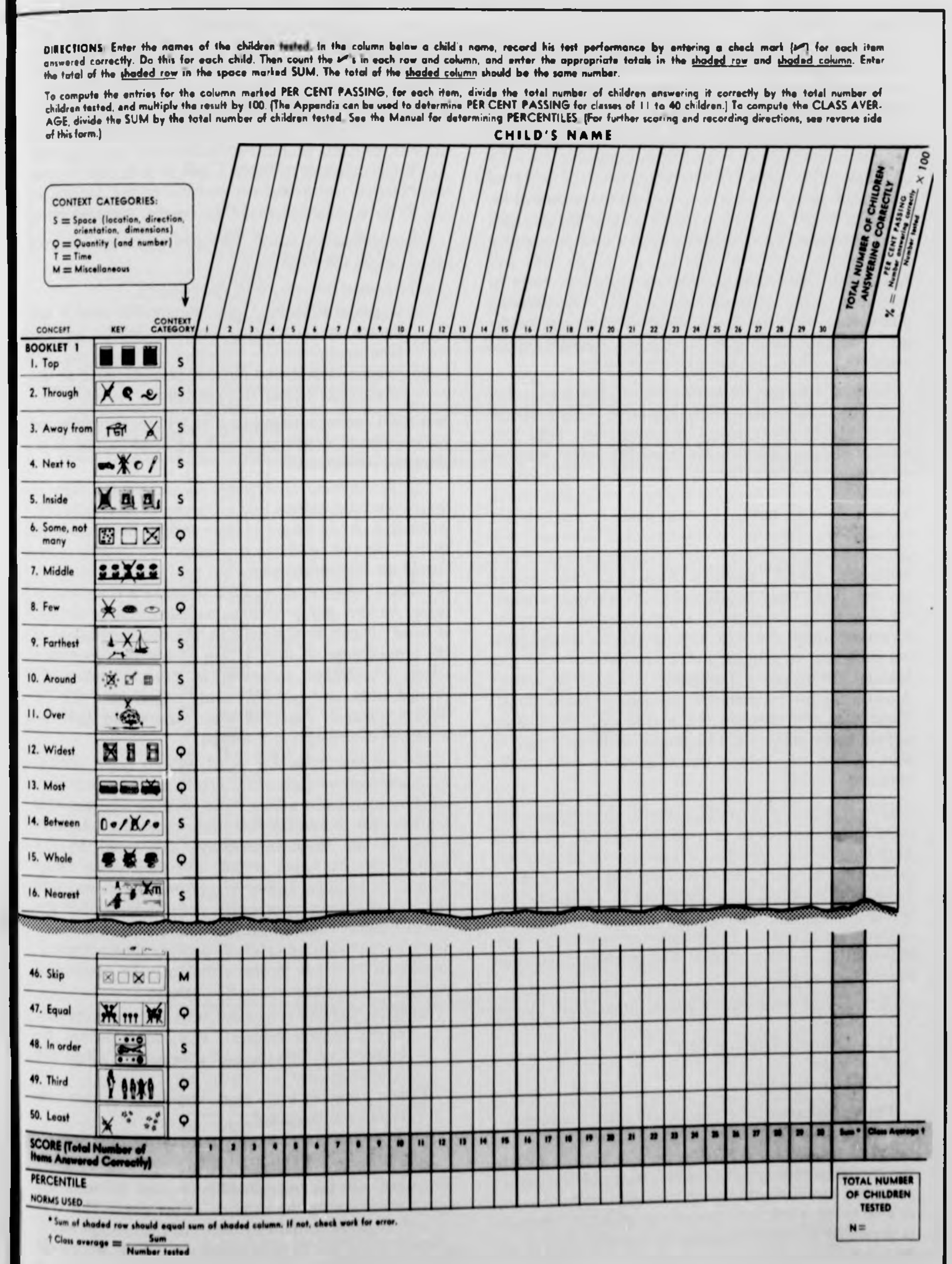
On the other hand, if the class average is high, intensive group instruction is probably unnecessary and attention need be given only to the occasional concepts that are missed by a large proportion of the class. Individual attention is needed only by those children who perform poorly. One way of selecting such children would be to identify all those who obtain scores below the class average. The degree to which remedial training of such children can be individualized depends on the number of children in the group thus selected.

To summarize, the teacher should base any choice of remedial instructional procedures upon two considerations: the average performance of the class as a whole, and the individual performance of specific children. When most of the children in a class obtain low scores, group training procedures are recommended. When the group as a whole performs well, special instruction for individual children who score below the class average is desirable.

Suggestions for Use of the Results in Remediation

The *BTBC* has been designed to assess children's knowledge of important concepts in their simple forms. With three exceptions, items are of a straightforward type; that is, different concepts have not been combined in the same item. In only one item, item 6, is the child asked to distinguish between similar concepts differing in degree. (Form A: "Mark the box that has *some but not many* marbles." And Form B: "Mark the bowl that has *some but not many* flowers.") None of the items require the child to distinguish between comparatives and superlatives, such as "bigger than one but not the

Class Record Form for Form A (partial reproduction, reduced from 24⅝" x 11").


DIRECTIONS: Enter the names of the children tested. In the column below a child's name, record his test performance by entering a check mark (✓) for each item answered correctly. Do this for each child. Then count the ✓'s in each row and column, and enter the appropriate totals in the shaded row and shaded column. Enter the total of the shaded row in the space marked SUM. The total of the shaded column should be the same number.

To compute the entries for the column marked PER CENT PASSING, for each item, divide the total number of children answering it correctly by the total number of children tested, and multiply the result by 100. (The Appendix can be used to determine PER CENT PASSING for classes of 11 to 40 children.) To compute the CLASS AVERAGE, divide the SUM by the total number of children tested. See the Manual for determining PERCENTILES. (For further scoring and recording directions, see reverse side of this form.)

CONTEXT CATEGORIES:
S = Space (location, direction, orientation, dimensions)
Q = Quantity (and number)
T = Time
M = Miscellaneous

CHILD'S NAME

CONCEPT	KEY	CONTEXT CATEGORY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	TOTAL NUMBER OF CHILDREN ANSWERING CORRECTLY	% = PER CENT PASSING Number answering correctly / Number tested × 100
BOOKLET 1 1. Top		S																																
2. Through		S																																
3. Away from		S																																
4. Next to		S																																
5. Inside		S																																
6. Some, not many		Q																																
7. Middle		S																																
8. Few		Q																																
9. Farthest		S																																
10. Around		S																																
11. Over		S																																
12. Widest		Q																																
13. Most		Q																																
14. Between		S																																
15. Whole		Q																																
16. Nearest		S																																
46. Skip		M																																
47. Equal		Q																																
48. In order		S																																
49. Third		Q																																
50. Least		Q																																
SCORE (Total Number of Items Answered Correctly)			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Sum*	Class Average†
PERCENTILE NORMS USED ____																																		

TOTAL NUMBER OF CHILDREN TESTED
N =

* Sum of shaded row should equal sum of shaded column. If not, check work for error.

† Class average = $\frac{\text{Sum}}{\text{Number tested}}$

biggest." Similarly, no item attempts to assess whether the child knows that the *top* of a ladder becomes its *bottom* if the ladder is turned upside down, or that *around* can mean "encircling" or "in the vicinity of," depending upon the context.

The items were prepared in this relatively simple form because it is in this way that the concepts they embody are most often used in the "Directions" portions of curriculum materials. For the same reason, remedial instruction should start with a presentation of the concepts in an elementary context before more subtle or complex connotations of the concepts are presented. This will help insure that lagging children assimilate minimal meanings of the concepts necessary for understanding instructions. Needless to say, if post-training test results are to be used for evaluative purposes, teachers should avoid the use of actual test items in training.

Selecting Concepts for Instruction.—A group remedial program should begin with identification of the concepts represented by items missed or omitted by a large percentage of the children in the class. The *BTBC* is a test of mastery of basic concepts, and hence any concept understood by less than 100 per cent of the children in a class should, ideally, be considered a problem for remediation. Practical considerations, however, will dictate the number of concepts that can be covered and hence the "cut-off" point, in terms of per cent passing an item, that should be adopted when selecting concepts for class study. As an aid in determining this point, information about the percentage of various groups passing each item is reported in the Norms section of this Manual. (Information is presented separately for Forms A and B and for beginning-of-year and midyear testings.) These data will acquaint the teacher with the characteristic performance of children in different types of schools, and will help in the selection of concepts for training.

Classification of Concepts.—Once the concepts for remediation have been selected, the teacher may wish to group them into "context categories," so that all those pertaining to, say, quantity can be discussed together. In the column immediately following the scoring key on the left side of the Class Record Form, a category describing the context in which each item is used has been indicated by a letter symbol. The categories and their symbols are as follows:

S = Space (location, direction, orientation, dimensions)
Q = Quantity (and number)
T = Time
M = Miscellaneous

The classification of concepts into these four context categories is arbitrary to some extent, and other meaningful classifications could doubtless be suggested. Still, the groupings reflected on the Class Record Form should provide a useful starting point for organizing the teacher's remedial work. A complete list of *BTBC* concepts is presented in Table 1 together with their classification by context category.

Reasons for Concept Difficulty.—Special difficulties may arise when a concept falls into more than one category, as in the use of the word *between* in "What is *between* the spoons?" (Space) and "What meal do we eat *between* morning and afternoon?" (Time). Knowledge of a concept in one context does not insure familiarity in another. Multiple classifications are suggested for certain concepts in Table 1, but in each case the one classification which best fits the use of the concept in the *BTBC* is emphasized for the teacher's convenience.

Other difficulties which may contribute to incorrect responses on the *BTBC* are:

(a) inherent difficulty with the concept,
(b) language difficulty where the specific label is not known or alternative concept labels are unfamiliar, or
(c) pictorial representation that is ambiguous or out of the child's realm of experience.

If a child cannot respond to a test item, therefore, the teacher should determine which of these factors is influencing performance.

The presentation of concepts on the printed page involves several problems that are more or less important, depending on the nature of the concept. Some concepts such as *big* and *small* can be clearly and directly represented on the printed page. No previous knowledge of the object presented is required to make a correct judgment. Another group of concepts, such as *near* and *far*, is more difficult to represent in the two dimensions of the printed page. A third group of concepts requires a degree of abstract judgment in comprehending the printed item, and possibly some experience with the objects pictured. *Fast* and *slow*, *heavy* and *light* are examples of concepts of this type.

Starting Instruction.—A good way to begin instruction is with an analysis of the errors made by individual children, starting by asking those children who answered a given item incorrectly how they chose their answers. Some of the misunderstandings uncovered by this method may be rather subtle, and their exact nature difficult to identify. However, personal interviews with children who participated in the pre-standardization tryout of Form A items yielded a number of typical errors. These are summarized below to provide the teacher with an idea of the kinds of errors that may be expected of children in kindergarten and first grade:

(a) Some children responded with the polar opposite of the concept tested. ("It's littlest because it's biggest," or "It's the widest because it's little and kind of skinny.")
(b) Some marked more than one choice so that confusion was suggested.
(c) Some responded to only part of the direction.
(d) Some had no notion of the concept and guessed.

Response with the opposite concept was the most frequent error among the children interviewed. This type of error suggests partial knowledge of the concept tested, and confusion due to poor establishment of the concept or the concept label.

Table 1. Classification of BTBC Concepts*

	Concept	Context Category: Space	Quantity	Time	Miscellaneous
1	Top	x			
2	Through	x			
3	Away from	x			
4	Next to	x		*	
5	Inside	x			
6	Some, not many		x		
7	Middle	x		*	
8	Few		x		
9	Farthest	x		*	
10	Around	x			
11	Over	x			
12	Widest		x		
13	Most		x		
14	Between	x		*	
15	Whole		x		
16	Nearest	x		*	
17	Second	*	x	*	
18	Corner	x			
19	Several		x		
20	Behind	x			
21	Row	x			
22	Different				x
23	After	*		x	
24	Almost		x		
25	Half		x		
26	Center	x			
27	As many		x		
28	Side	x			
29	Beginning	*		x	
30	Other				x
31	Alike				x
32	Not first or last	*	x	*	
33	Never			x	
34	Below	x			
35	Matches				x
36	Always			x	
37	Medium-sized		x		
38	Right	x			
39	Forward	x			
40	Zero		x		
41	Above	x			
42	Every		x		
43	Separated	x		*	
44	Left	x			
45	Pair		x		
46	Skip				x
47	Equal		x		
48	In order	x			
49	Third	*	x	*	
50	Least		x		

* X's indicate the context category of each concept as it is tested by the *BTBC*; asterisks indicate additional contexts in which the concepts may be employed. For example, the concept of *beginning* (item 29) is used in the context of time on the *BTBC*, but it may also be used to express relationships involving space.

Children who answered correctly were generally able to do one or more of the following:

(a) Describe the question using a synonymous concept label. ("It's thickest because it's biggest.")
(b) Eliminate choices within an item as incorrect. ("It wasn't longest or the little one.")
(c) Specifically point out cues used in pictures. ("It's in front of the dog because his head is looking at it.")

Using the above material as general background, the teacher may move naturally from discussions of error analyses into activities designed to increase the *depth* and *scope* of the children's understanding. Thus, the several meanings and connotations of a concept may be explored, as well as its use in different situations and contexts.

Sequence of Concept Presentation.—The teacher may find it useful to build a regular sequence into the presentation of basic concepts. This sequence might include:

(a) presenting the concept through the use of concrete materials,
(b) specifically labeling the concept,
(c) using the concept in several concrete situations so that the label does not become tied to one object or situation,
(d) showing a representation of the concept in photographs or simple pictures, and
(e) finally using the concept in a tailor-made workbook format (being sure to have the pictures large and well-spaced on the page).

As children gain a firmer grasp of basic concepts, the teacher can introduce concepts that are used in different ways and that fall into more than one category. For example, the children can be taught that *through* has the meaning of "finished" as well as its directional meaning; they can be shown the various applications of the concept *corner;* and they can discover that an object can be both *short* and *long,* depending upon the reference object with which it is compared. (The multiple classification of items by context category in Table 1 will help the teacher identify concepts with several meanings.)

The depth and scope of concept usage can be fostered by:

(a) using concepts in *combination* (for example, "Point to the *last* box in the *first* row."),
(b) focusing attention on *opposites* (for example, *left—right, top—bottom*),
(c) focusing attention on *degree* (for example, *far, farther, farthest*), or
(d) expressing concepts through more than one *sense modality* (for example, a *high table* and a *high musical note,* in the visual and auditory modalities, respectively).

Pre- and Post-Training Testing.—An especially important use of the *BTBC* is as a post-training measure of the mastery of basic concepts. Initial results on the *BTBC* provide a sound basis for planning the content and sequence of training. Administration of the alternate form of the test upon completion of training then helps to indicate the progress of the children in the program or, viewed differently, the effectiveness of the remedial work. Form B was designed to minimize the effects of practice on scores obtained on retesting. As noted earlier, training should not focus on the specific items on either form of the test, if post-training test results are to reflect more than rote memorization.

Although Forms A and B yield equivalent total scores, the difficulty levels for most corresponding items on the two forms (those that measure the same concepts) do differ slightly, and a few items are quite dissimilar in difficulty. (For a detailed discussion of the equivalence of Forms A and B, see the Technical Data section of this Manual.) Such differences need to be taken into account when interpreting the results of such studies that employ one form as a pre-training measure and the alternate form as a post-training measure.[7]

Despite these differences, the administration of alternate forms of the test as pre- and post-training measures is preferable to the repeated use of the same form, since children are thereby allowed to demonstrate their concept knowledge in more than one situation. In this regard the following points should be kept in mind:

1. Before-and-after comparisons of scores for an individual student are often difficult to interpret—indeed, they are sometimes hard to justify—because of the statistical unreliability of difference scores. This is especially true of scores on individual items. Consequently, the difference between pre-training and post-training scores of a particular pupil should be interpreted tentatively and with caution.

Differences between *a group's average performance* on an item before and after training are statistically more stable than an individual's item scores, and may be used with some confidence to demonstrate group changes. These differences may reflect the overall effectiveness of a concept-training program, and provide evidence relevant to program evaluation.

Item scores are informative about children's functioning. In considering an upward shift in item scores, several possible sources of change may be operating

[7] An example of an item whose difficulty differs markedly between the two forms would be item 39, which measures the concept "forward." (Form A: "Mark the boy who is bending *forward*." And Form B: "Mark the chick that is bending *forward*.") Table 5 indicates that for the beginning-of-year norms group, 33 per cent of the kindergarten children at the low-socioeconomic-level schools answered the item correctly on Form A. However, Table 7 shows that 83 per cent of the children at the same grade and socioeconomic level answer the parallel item on Form B correctly. Clearly the Form B item is easier than the corresponding Form A item. Thus if Form A were administered before training and Form B after training, a simple increase in the percentage of children answering item 39 correctly would not necessarily indicate success in teaching students the concept "forward." Similarly, if the two forms were administered in reverse order, a decrease in the percentage of children answering item 39 correctly would not necessarily indicate that the training program had been detrimental to the children's conceptual knowledge!

(such as improved concept understanding, better test-taking skills, or increased maturation on the part of the children).

2. Post-training scores are interpretable in themselves, without reference to pre-training results. For individual pupils they reflect current status and may be used in the same way as pre-training scores. Group summaries—particularly difficulty values for individual items—may be used to identify concepts on which a class could profit from further instruction.

3. The effects of different item difficulty values across the forms can be offset to some extent if half of the children in a given group receive Form A as a pre-training measure and Form B as a post-training measure, with the remainder of the children taking the two forms under similar conditions but in the reverse sequence. Such a procedure will be facilitated if two comparable classroom groups of about equal size are available. Then pre-training difficulty values for each of the 50 concepts may be computed by pooling the results obtained with Forms A and B given before training, and post-training difficulty values may be similarly obtained.

Through this procedure the value of using alternate forms is retained, since no child is tested twice with the same form of the *BTBC*. Also the effects of different item difficulties on Forms A and B are canceled out, since the obtained percentages for the total number of students demonstrating mastery of a given concept are based on the administration of both forms at each testing. Thus, within the bounds imposed by error of measurement, the pre- and post-training differences in the percentages of students demonstrating knowledge of a concept would reflect only the effects of training and of maturation.

4. The difference in the difficulty of corresponding items on the two forms is less pronounced for groups at higher grade and socioeconomic levels. Item 39, which was cited earlier as an extreme example of differential difficulty across the two forms, may be used here for illustration. For the beginning-of-year norms, the difference in item difficulty for the item versions appearing on the two forms is .50 for the kindergarten groups at low-socioeconomic-level schools. (The difficulty value for the Form A version of the item is .33, and for the Form B version it is .83.) For the grade 2, high-socioeconomic-level norms group, the difficulty value of the same item is identical for both forms (.93). Generally, these discrepancies between difficulty values of corresponding items decrease at progressively higher grade and socioeconomic levels. Thus differences in item difficulty will have less effect on the test results of more able groups of children.

Concepts Omitted from the BTBC.—For the benefit of teachers whose pupils have had extreme difficulty with the test, a list of the concepts *not* included in the test is presented in Table 2. These concepts were drawn from the same sources as those included in the test, but were eliminated on the basis of the tryout results. These results showed, it will be remembered, that most of the children in the samples studied did know the concepts *in the forms in which they were tested*—these forms having been chosen because they reflected the level of meaning characteristic of instructional materials for kindergarten and first grade. This does not mean, however, that children at these grade levels can be expected to possess "depth knowledge" of the concepts in question. It is possible and, in fact, probable that some children have been so severely deprived of concept-learning experiences that they will find even these "familiar" concepts too difficult when presented in printed form.

Hence, teachers who undertake an enrichment program of the type described earlier will not necessarily wish to restrict themselves to the concepts contained in the test itself. Table 2 has been provided for the use of such teachers. It is worth noting that the concepts listed in Table 2, together with those included in the *BTBC*, constitute virtually the entire set that fulfilled the item-selection criteria used during the initial review of curriculum materials.

Pre-Kindergarten Use of the BTBC

Teachers at the pre-kindergarten level may wish to adapt the *BTBC* for use with their classes. This might be done, for example, by administering the test individually to each child, and requesting the child to *point to* the answer instead of marking it with a crayon. In using test results with pre-kindergarten groups, it is important to keep in mind that some of the basic concepts measured are very difficult for the average four-year-old child to master, particularly those in the latter portion of the test.

The above suggestions for the use of *BTBC* results in remedial and enrichment teaching are intended to illustrate a few of the educational possibilities of the test. The imaginative teacher will find many ways to supplement these suggestions and to incorporate them into a teaching program especially designed to meet local needs.

Table 2. Classification of Concepts Not on BTBC

Concept	Context Category: Space	Quantity	Time	Miscellaneous
Bottom	x			
Under	x			
Beside	x			
In front	x			
Toward—away	x			
Beyond	x			
Up—down	x			
Big—small	x	x		
Tall—short	x	x		
Long—short	x	x	x	
Wide, fat—narrow, thin	x	x		
Round	x			
Flat	x			
Straight	x			
Line	x			
Follow	x		x	
Join	x			
Change				x
Moving—still	x			
Deep—shallow	x			
Fast—slow			x	
Now			x	
Early—late			x	
Past			x	
Start, begin—stop, finish			x	
Open—closed				x
Soft—hard				x
Easy—hard				x
Dark—light				x
Loud				x
Light—heavy		x		x
Any		x		
Every—none		x		
Enough		x		
With—without				x

TECHNICAL DATA

Empirical Item Selection Procedures

The criteria governing the selection of concepts for inclusion in the *BTBC* were described earlier in this Manual. Briefly, these criteria required that the concepts selected be important for understanding and following instructions, occur frequently in curriculum materials, and have little or no attention given to their instruction. Application of these criteria resulted in the selection of a set of basic concepts that authors of preschool and primary instructional materials frequently assume children do know at the time of their entry into the school system.

Tryout Samples for Form A

The purpose of the experimental tryout of items was to identify those that would be most useful to the teacher in terms of the test's rationale. The first tryout sample consisted of 267 kindergarten, 306 first-grade, 264

second-grade, and 297 third-grade pupils attending four schools in a city of about 75,000 population located in southeastern New York State. One of the participating schools drew its pupils from population areas of relatively high socioeconomic status, one drew from areas of relatively low socioeconomic status, and two drew from mixed areas.

During October 1967, eight 25-item test booklets containing, in all, 200 tryout items were administered to the children at each grade level. Since at least two items had been prepared for each concept, the booklets covered a total of approximately 100 concepts.

Due to time limitations each class of pupils was given only four of the eight booklets, with the selection and sequence of booklets rotated to help assure random tryout samples for each item.

The analysis of the data consisted of determining, separately for each of the four grade levels sampled, the proportion of the children who passed each item, and the point-biserial coefficient of correlation of each item with the subjects' total scores on the booklet in which the item appeared. The best 50 items were selected on the basis of these statistics. These items (a) each measured a different concept, (b) had point-biserial correlations exceeding .30, (c) showed fairly even rises in per-cent-passing values across age levels, and (d) together yielded a roughly normal distribution of per-cent-passing values, centered around .50, for the kindergarten pupils.

After minor revisions intended to increase brevity and clarity, the 50 items were arranged in approximately ascending order of difficulty in two 25-item booklets, and administered to a second sample of children. This sample consisted of 671 kindergarten, 823 first-grade, and 710 second-grade pupils from six schools in a city of about two million population located in an eastern state. Third-grade children were not included; they found the items so easy in the first tryout that the usefulness of the test for average children at this grade level is questionable.

On the basis of the data obtained from the second tryout, a few additional item revisions were made. The items were then rearranged to improve their order of difficulty. The resulting revision of the *BTBC* was used to secure normative data.

Standardization of Form A

The *BTBC* was designed as a screening and teaching instrument rather than for predictive or administrative purposes. Consequently, it was considered unnecessary to select standardization samples representative of children in kindergarten and the first and second grades in the nation as a whole. Since *BTBC* scores may be interpreted on an absolute or a relative basis, the normative information presented in this Manual is intended more as an informal guide than as the essential procedure for interpreting test results.

The Form A standardization sample which served as the basis for *beginning-of-year* norms consisted of children enrolled in kindergarten and the first and second grades in each of sixteen cities located across the United States. Testing was conducted in September and October of the 1969–70 school year. Data for *midyear* norms were obtained during the 1968–69 school year, from testing conducted between mid-November and late February. The sample included children from schools in five cities.

School officials in each cooperating city were asked to provide classroom groups from schools with a fairly wide range of socioeconomic background. Selection was to be based on the socioeconomic level of the *primary areas* from which the schools drew their enrollments. No formal specifications for this selection were given; the choice was left to the judgment of the administrative personnel. Table 3 presents the number of pupils tested on Form A in each city, by grade and socioeconomic level, for both the beginning-of-year and the midyear standardization samples.

Selection of Content for Form B

Each item of Form B was designed to measure knowledge of the same concept as that measured by the corresponding item of Form A. For example, item 1 on both forms measures mastery of the concept "top."

One difficulty in constructing parallel items for a test of concept information results from the very nature of conceptual knowledge. A concept is a general notion, an abstraction applicable in many contexts. The concept "beginning," for instance, could pertain to the start of the day, the commencement of an action, the first of a series of objects, or countless other referents, and yet retain its essential meaning. In writing items for Form B an attempt was made to present each item in a context similar to that of the Form A item. However, minor differences which did not seem to alter the essential nature of a concept's usage were utilized. An "armchair" attempt was also made to match the difficulty of Form B items to that of the corresponding items in Form A. Thus the two forms of the *BTBC* were designed to parallel each other in both coverage and difficulty level.

Equivalence of Forms A and B

In order to study the equivalence of scores on Forms A and B, both forms of the test were administered to groups of children enrolled in kindergarten and the first and second grades in schools located in three different cities. The testing was conducted in December and January of the 1970–71 school year.

As with the standardization of Form A, the sample of each grade was subdivided by socioeconomic level, based on the judgments of the local school administrators. Approximately half of the children at each socioeconomic level in each grade were administered Form A first, followed by Form B given on a different day but within one week of the first testing. The other half of the children were tested at the same time, but with the two forms administered in reverse order. Table 4 presents the number of pupils in each city, by grade and socioeconomic level, participating in the first testing

Table 3. Standardization Samples for Form A by City, Grade, and Socioeconomic Level

City	Kindergarten — Socioeconomic Level: Low	Middle	High	Total	Grade 1 — Socioeconomic Level: Low	Middle	High	Total	Grade 2 — Socioeconomic Level: Low	Middle	High	Total
Beginning of Year:												
Anaheim, Calif.	—	47	36	83	—	78	72	150	—	41	72	113
Los Angeles, Calif.	—	—	18	18	513	214	246	973	—	—	—	—
San Francisco, Calif.	30	—	43	73	—	—	64	64	—	—	—	—
Washington, D. C.	75	—	60	135	87	—	60	147	118	—	60	178
Clearwater, Fla.	115	139	—	254	229	309	—	538	—	—	—	—
Hutchinson, Kans.	174	40	158	372	176	39	145	360	167	50	173	390
Minneapolis, Minn.	84	103	85	272	—	—	—	—	—	—	—	—
Columbia, Mo.	41	30	—	71	219	60	—	279	43	—	—	43
St. Louis, Mo.	—	—	—	—	219	182	—	401	—	—	—	—
Sedalia, Mo.	—	—	—	—	52	58	—	110	60	51	—	111
Cincinnati, Ohio	71	—	74	145	61	—	75	136	—	—	—	—
Philadelphia, Pa.	106	226	—	332	137	176	—	313	157	162	—	319
Dallas, Tex.	396	30	—	426	245	—	213	458	—	—	—	—
Norfolk, Va.	170	—	—	170	138	72	74	284	202	77	51	330
Richmond, Va.	545	149	98	792	56	—	—	56	77	—	—	77
Seattle, Wash.	114	148	112	374	171	125	94	390	—	—	—	—
Total	1921	912	684	3517	2303	1313	1043	4659	824	381	356	1561
Midyear:												
Fresno, Calif.	22	58	61	141	68	73	71	212	—	60	34	94
Atlanta, Ga.	63	43	26	132	87	77	38	202	85	71	27	183
Highland Park, N. J.	—	125	48	173	—	74	45	119	—	62	61	123
New Rochelle, N. Y.	12	154	36	202	67	103	46	216	61	134	40	235
Tulsa, Okla.	65	73	79	217	54	108	80	242	76	22	80	178
Total	162	453	250	865	276	435	280	991	222	349	242	813

Table 4. Sample for Equivalence Study of Forms A and B, by City, Grade, and Socioeconomic Level[a]

City	Kindergarten — Socioeconomic Level: Low	Middle	High	Total	Grade 1 — Socioeconomic Level: Low	Middle	High	Total	Grade 2 — Socioeconomic Level: Low	Middle	High	Total
Billings, Mont.	—	54	—	54	46	63	36	145	38	52	44	134
Binghamton, N.Y.	45	36	43	124	43	42	41	126	48	44	52[b]	144
Philadelphia, Pa.	41	52	49	142	38	65	62	165	39	58	61	158
Total	86	142	92	320	127	170	139	436	125	154	157	436

[a] The cases summarized here are the pupils participating in the first testing sessions of the equivalence study, regardless of the form of the test that was administered. These groups were used for studying the equivalence of *total scores* on Forms A and B. A few pupils were absent for the first testing session of the equivalence study, but did take part in the second session; these cases were added to the above sample for establishing the equivalence of *corresponding items* on Forms A and B.

[b] These pupils are from two classroom groups, both of which were inadvertently administered the forms of the *BTBC* in the same sequence (Form A followed by Form B). For this reason the scores of these groups were excluded from all data analyses except the computation of alternate form reliability coefficients.

session, regardless of whether they took part in the second session.

Norms for total scores and for individual items on Form B were developed from data gathered in this manner. Details of the procedure are given below.

Equivalence of Total Scores

Only results of the first testings of the 1970–71 sample were employed to study the equivalence of total scores on the two forms of the *BTBC*. Data for Forms A and B, based on separate but comparable samples (N = 570 for each form), were available for study.[8] The mean scores on Forms A and B were 42.4 and 42.9, respectively; the standard deviations of these scores were 7.3 and 7.0. These figures reflect the overall comparability of scores on the two forms.

It is also important to study the comparability of the forms throughout the total range of scores, and to this end the forms were equated through an equi-percentile procedure. Frequency distributions of the raw scores were prepared separately for the two forms, and percentile equivalents were computed for each raw score. A graph was prepared for each form, showing the relationship between each raw score and its corresponding percentile. Perfect equivalence of the forms would have been indicated had the two graphs overlapped throughout their ranges. While the graphs did not correspond perfectly, at no point were scores on Form B more than 1.5 points higher than the comparable scores on Form A. Thus their degree of overlap permitted the conclusion that the two forms yield essentially equivalent raw scores. The difference between mean scores on the two forms, as indicated above, was only .5.

Equivalence of Scores on Corresponding Items

The same sample of pupils used to study the equivalence of total raw scores obtained on Forms A and B was used to establish the equivalence of corresponding items on the two forms. Data for both the first and the second testing sessions were pooled, however, for this phase of the analysis. For Forms A and B separately, the data were divided into nine subsamples by grade (kindergarten and grades 1 and 2) and socioeconomic level (low, middle, and high). Next, for each of the 50 items on Form A, a difficulty (or per-cent-passing) value was obtained for each of these subsamples; nine difficulty values were similarly computed for each of the corresponding items on Form B.

Based on these values, a separate regression equation was established for each concept, so that the difficulty value of a Form B item could be estimated for a specified group of children, given a knowledge of that group's performance on the corresponding Form A item. These equations were then used for deriving norms tables for the Form B items, as described in the Norms section below.

[8] The total number of cases used for this portion of the analysis is less than would be expected from the N's reported in Table 4, because the two second-grade, high-socioeconomic-level classroom groups in one city were both inadvertently tested with Form A at the first session and Form B at the second session; data from these groups were omitted from this analysis.

Norms

The per cent of children passing each individual item on the *BTBC* will be of greatest interest to the kindergarten and primary-grade teacher. Table 5 presents these percentage figures for Form A items for the beginning-of-year standardization sample separately by grade, and by socioeconomic level within each grade. Table 6 presents corresponding Form A data for the midyear sample. As expected, both sets of data show that the concepts become better known with increasing grade level, and also with increasing socioeconomic status within each grade. In addition, the midyear data reflect a generally greater mastery of concepts than is found at the beginning of the year; for most items, the percentage figures in Table 6 are higher than the corresponding percentages given in Table 5. The exceptions occur mainly at the middle and high socioeconomic levels of grade 2, where the items tend to be quite easy regardless of the month of testing.

Tables 7 and 8 present, for beginning-of-year and midyear testings, respectively, the estimated per cent of children passing the Form B items, based on the application of the regression equations described above to the Form A difficulty values given in Tables 5 and 6.

An illustration of the use of the regression equations is given here to clarify the procedure. The standardization data for Form A indicated that 58 per cent of the kindergarten pupils in schools drawing their enrollment from areas of low socioeconomic level passed item 1 (the concept of *top*) at the beginning of the school year. The regression equation for this item showed that an estimated 64 per cent of the same group would pass the corresponding item on Form B. Similarly, 97 per cent of the standardization sample of pupils in grade 1 at middle-socioeconomic-level schools passed the same item at the beginning of the year, and the regression equation for this item indicates that 96 per cent of these children would be expected to pass the corresponding Form B item. Thus the difficulty values given in the norms tables for Form A were entered into the regression equations to obtain estimates of Form B norms for the original standardization sample.

A study of the corresponding per-cent-passing values for Forms A and B indicates that most of the items on the two forms are nearly equivalent. In some cases, however, one of a pair of items yields substantially higher difficulty values than does the corresponding item on the other form. For such item pairs, care should be taken in comparing the performances of groups which took different forms or in comparing the performances of a single group which took different forms on separate occasions—for example, at the beginning and end of the school year. The apparent difference in Form A and Form B item difficulties cannot be discounted in interpreting results of testing with the two forms. Some of the issues involved in comparing the

(*Text continued on page 26.*)

FORM A
BEGINNING OF YEAR

Table 5. Per Cent Passing Each Item, by Grade and Socioeconomic Level
(N=9737 Children Tested at Beginning of School Year)

Item	Kindergarten Socioeconomic Level: Low	Middle	High	Grade 1 Socioeconomic Level: Low	Middle	High	Grade 2 Socioeconomic Level: Low	Middle	High
Booklet 1									
1	58	67	73	85	97	98	97	99	96
2	77	85	92	91	98	99	98	99	99
3	57	69	77	84	96	97	97	100	99
4	64	82	88	86	98	98	96	99	99
5	77	90	94	93	99	100	98	100	100
6	74	84	91	90	98	98	97	99	99
7	70	87	92	89	99	99	94	98	99
8	63	84	90	73	88	91	83	94	96
9	63	87	93	86	98	99	95	100	100
10	82	91	94	94	97	99	98	100	100
11	71	84	90	81	96	99	92	99	100
12	66	81	83	79	95	96	95	98	100
13	79	84	90	89	96	98	97	99	100
14	64	80	84	77	92	92	89	98	98
15	69	81	84	79	92	92	91	99	97
16	90	96	96	96	99	100	98	100	100
17	47	60	70	69	91	92	95	99	99
18	64	78	86	83	96	96	95	99	99
19	75	83	87	84	90	93	90	97	98
20	71	87	89	80	94	96	93	99	99
21	49	72	83	80	95	97	94	99	100
22	53	73	86	80	95	98	93	99	100
23	60	72	82	80	92	97	91	96	97
24	65	69	72	76	83	88	90	95	96
25	68	75	81	82	88	88	94	97	99
Booklet 2									
26	60	68	70	64	86	90	78	93	99
27	29	52	63	62	90	92	84	99	97
28	45	55	65	64	84	85	81	88	92
29	48	68	76	68	85	89	87	94	94
30	43	66	78	71	88	90	86	98	97
31	45	54	64	63	82	88	83	94	95
32	40	60	74	58	85	92	84	99	98
33	32	44	66	57	78	83	84	94	96
34	34	60	82	58	89	92	76	98	99
35	45	53	63	60	81	82	77	90	90
36	31	53	62	55	80	85	74	93	94
37	29	47	65	40	69	80	52	89	89
38	35	42	50	50	65	70	66	81	89
39	33	52	61	40	72	82	59	89	93
40	19	32	44	52	78	85	89	98	99
41	45	65	78	59	87	92	76	95	98
42	69	77	89	77	92	95	89	99	98
43	39	45	51	53	64	73	67	91	87
44	36	45	53	56	70	75	63	82	87
45	28	21	24	29	39	41	33	67	62
46	17	28	41	45	68	77	80	94	94
47	14	16	18	26	40	45	42	71	67
48	27	31	42	32	60	73	59	91	92
49	16	27	33	36	65	70	67	83	81
50	18	16	19	22	36	43	33	64	73
N	1921	912	684	2303	1313	1043	824	381	356

FORM A
MIDYEAR

Table 6. Per Cent Passing Each Item, by Grade and Socioeconomic Level
(N=2669 Children Tested at Midyear)

Item	Kindergarten Socioeconomic Level: Low	Middle	High	Grade 1 Socioeconomic Level: Low	Middle	High	Grade 2 Socioeconomic Level: Low	Middle	High
Booklet 1									
1	62	82	89	99	99	98	98	99	100
2	86	91	97	99	100	99	100	100	100
3	59	81	92	96	99	99	99	100	100
4	73	92	92	98	100	100	98	99	100
5	82	93	97	99	100	100	100	100	100
6	84	91	93	98	99	99	99	100	100
7	80	93	96	96	98	99	95	100	99
8	60	90	92	87	90	93	90	95	97
9	71	92	96	92	99	99	96	99	100
10	83	94	99	100	100	100	100	100	100
11	72	91	95	86	96	99	98	99	100
12	70	83	88	92	97	99	95	100	100
13	85	92	94	96	99	99	99	99	100
14	68	85	86	90	92	94	91	97	98
15	65	82	84	84	90	95	94	98	99
16	93	97	98	99	100	100	97	100	100
17	52	71	81	87	94	96	98	99	98
18	71	84	88	91	98	99	97	100	99
19	81	86	91	85	91	95	89	97	97
20	74	88	91	92	98	97	98	99	98
21	59	84	92	92	99	99	95	99	99
22	69	86	91	94	100	98	97	100	100
23	72	83	87	85	92	94	91	94	98
24	75	79	77	84	88	95	86	96	96
25	74	78	82	87	92	94	92	97	97
Booklet 2									
26	64	71	79	76	87	92	75	93	94
27	34	58	76	81	93	98	89	97	99
28	48	65	71	75	82	87	82	89	88
29	59	73	86	77	90	95	90	95	95
30	58	70	91	84	95	95	91	98	99
31	60	56	76	80	92	95	91	93	96
32	49	72	87	79	93[a]	96	93	100	98
33	38	61	75	75	90	91	92	96	99
34	51	76	86	87	95	95	88	97	99
35	46	55	63	61	83	81	69	81	90
36	34	66	81	68	83	93	84	96	97
37	20	52	63	41	68	82	56	82	91
38	45	46	63	67	74	79	82	85	90
39	45	57	68	51	80	89	75	91	95
40	24	40	60	83	97	99	96	99	100
41	55	69	80	70	88	92	86	96	98
42	71	86	90	83	93	96	93	97	99
43	49	47	64	63	76	82	75	90	94
44	43	49	63	61	78	85	78	85	91
45	26	32	37	36	40	50	43	60	73
46	27	42	57	70	83	90	85	96	95
47	10	11	33	34	49	50	63	66	71
48	26	39	49	42	71	82	71	87	95
49	21	39	44	47	69	75	68	76	78
50	18	22	31	29	49	51	44	64	81
N	162	453	250	276	435[a]	280	222	349	242

[a] Because item 32 appeared to have been administered incorrectly to one first-grade classroom group in a middle-socioeconomic-level school, all responses from that group were excluded when calculating the per cent passing item 32. For this item only, the N is 413 rather than 435.

FORM B
BEGINNING OF YEAR

Table 7. Per Cent Passing Each Item, by Grade and Socioeconomic Level*

Item	Kindergarten Socioeconomic Level: Low	Middle	High	Grade 1 Socioeconomic Level: Low	Middle	High	Grade 2 Socioeconomic Level: Low	Middle	High
Booklet 1									
1	64	71	76	86	96	97	96	98	95
2	86	90	94	93	97	98	97	98	98
3	90	93	95	96	99	99	99	100	100
4	79	89	93	92	98	98	97	99	99
5	81	92	95	94	99	100	99	100	100
6	40	58	71	70	84	84	83	86	86
7	76	89	93	91	98	98	95	98	98
8	72	87	91	79	90	92	86	94	95
9	32	74	84	72	93	95	88	97	97
10	87	93	95	95	97	98	98	99	99
11	63	78	85	75	92	95	87	95	96
12	82	89	90	88	96	96	96	97	98
13	78	83	89	88	96	98	97	99	100
14	73	86	89	83	96	96	93	100	100
15	59	74	78	71	87	87	86	96	93
16	86	94	94	94	98	99	97	99	99
17	50	62	71	70	90	91	93	97	97
18	86	91	94	93	97	97	97	98	98
19	77	84	88	85	91	93	91	97	98
20	83	93	94	89	97	98	96	100	100
21	43	67	79	76	92	94	91	96	97
22	16	50	73	62	88	93	85	95	97
23	54	68	79	77	91	97	90	95	97
24	81	83	85	87	90	93	94	97	97
25	15	31	44	47	60	60	74	81	86
Booklet 2									
26	69	75	76	72	87	89	81	91	95
27	55	67	73	72	87	88	84	92	91
28	95	96	97	96	99	99	98	99	99
29	67	81	86	81	92	95	93	98	98
30	51	71	82	75	90	92	89	99	98
31	50	58	67	66	83	88	83	93	94
32	55	70	81	69	89	94	88	99	98
33	14	26	49	40	62	67	68	79	81
34	64	78	89	77	93	94	86	97	98
35	54	61	70	67	86	86	82	94	94
36	95	97	97	97	99	99	98	100	100
37	38	53	68	47	71	80	57	87	87
38	51	55	60	60	70	73	71	80	86
39	83	86	88	84	90	91	87	92	93
40	32	43	53	60	82	88	91	99	100
41	62	76	86	72	92	96	84	98	100
42	65	74	87	74	90	94	87	98	97
43	41	47	54	56	67	76	70	95	91
44	44	51	58	60	71	76	66	81	85
45	33	26	29	33	42	44	37	67	62
46	16	27	40	44	67	77	80	94	94
47	15	17	19	27	42	47	44	73	69
48	49	51	59	52	71	80	71	92	93
49	14	25	31	34	64	69	66	82	80
50	24	22	25	28	43	50	40	72	81

* Percentage values estimated from corresponding data for Form A, based on results of equivalence study of Forms A and B. See Technical Data section of Manual for details of study.

FORM B
MIDYEAR

Table 8. Per Cent Passing Each Item, by Grade and Socioeconomic Level*

Item	Kindergarten Socioeconomic Level: Low	Middle	High	Grade 1 Socioeconomic Level: Low	Middle	High	Grade 2 Socioeconomic Level: Low	Middle	High
Booklet 1									
1	67	84	89	98	98	97	97	98	98
2	91	93	97	98	98	98	98	98	98
3	91	96	98	99	100	100	100	100	100
4	84	95	95	98	100	100	98	99	100
5	85	94	98	99	100	100	100	100	100
6	58	71	75	84	86	86	86	88	88
7	84	94	96	96	98	98	95	99	98
8	70	91	93	89	91	93	91	95	96
9	46	83	90	83	95	95	90	95	97
10	87	95	98	99	99	99	99	99	99
11	64	86	91	80	92	95	94	95	96
12	84	90	93	94	97	98	96	98	98
13	84	92	94	96	99	99	99	99	100
14	76	90	91	94	96	97	95	100	100
15	54	75	78	78	85	91	90	95	96
16	90	95	97	98	99	99	95	99	99
17	55	72	81	86	92	94	96	97	96
18	89	93	95	96	98	98	98	99	98
19	83	87	91	86	91	95	90	97	97
20	85	93	95	96	99	99	99	100	99
21	54	80	89	89	96	96	92	96	96
22	44	73	81	86	97	93	91	97	97
23	68	80	85	83	91	93	90	93	98
24	86	88	87	91	93	97	92	97	97
25	28	37	47	58	70	74	70	81	81
Booklet 2									
26	72	77	82	80	87	91	79	91	92
27	58	70	80	82	89	91	87	91	92
28	95	97	97	98	98	99	98	99	99
29	75	84	93	87	95	99	95	99	99
30	64	75	93	87	97	97	93	99	100
31	63	59	77	81	92	94	91	92	95
32	62	79	90	84	95	97	95	100	98
33	20	44	59	59	74	75	76	81	84
34	73	86	91	92	96	96	92	97	98
35	55	62	70	68	87	86	75	86	94
36	95	98	99	98	99	100	99	100	100
37	31	57	66	48	70	82	60	82	89
38	57	58	69	71	76	79	81	83	86
39	85	87	89	86	91	92	90	93	93
40	36	50	67	86	98	100	97	100	100
41	69	79	87	80	93	96	92	99	100
42	67	84	88	80	91	95	91	96	98
43	51	49	67	66	80	86	78	94	98
44	49	54	66	64	78	84	78	84	89
45	31	36	40	40	43	52	46	61	72
46	26	41	56	69	83	90	85	96	95
47	11	12	34	35	51	52	65	68	73
48	48	57	64	59	79	86	79	90	95
49	19	37	42	45	68	74	67	75	77
50	24	28	38	35	56	58	51	72	89

* Percentage values estimated from corresponding data for Form A, based on results of equivalence study of Forms A and B. See Technical Data section of Manual for details of study.

Table 9. Percentile Equivalents of Raw Scores, by Grade and Socioeconomic Level[a]
(N=9737 Children Tested at Beginning of School Year)

Percentile	Kindergarten, Socioeconomic Level: Low	Middle	High	Grade 1, Socioeconomic Level: Low	Middle	High	Grade 2, Socioeconomic Level: Low	Middle	High	Percentile
99	45-50	47-50	48-50	49-50	50	50	50	—	—	99
97	42-44	45-46	47	47-48	49	—	49	—	—	97
95	39-41	43-44	46	45-46	48	49	48	50	50	95
90	36-38	41-42	45	44	—	—	47	—	—	90
85	34-35	40	44	42-43	47	—	—	—	—	85
80	33	39	43	41	46	48	46	—	—	80
75	31-32	38	42	40	—	—	—	49	49	75
70	30	37	41	39	45	47	45	—	—	70
65	29	36	40	38	—	46	—	—	—	65
60	27-28	35	39	37	44	—	44	48	—	60
55	26	34	38	36	—	45	43	—	48	55
50	25	32-33	37	35	43	—	42	—	—	50
45	24	31	36	34	42	44	—	47	—	45
40	23	30	34-35	32-33	41	—	41	—	47	40
35	21-22	29	33	31	—	43	40	—	—	35
30	20	27-28	32	30	40	42	39	46	—	30
25	18-19	26	30-31	28-29	39	41	37-38	—	46	25
20	17	24-25	29	26-27	38	40	36	45	—	20
15	15-16	22-23	27-28	24-25	36-37	39	35	44	45	15
10	13-14	19-21	24-26	20-23	34-35	37-38	33-34	43	44	10
5	10-12	15-18	19-23	15-19	30-33	33-36	28-32	41-42	42-43	5
3	7-9	10-14	16-18	11-14	27-29	28-32	21-27	37-40	40-41	3
1	0-6	0-9	0-15	0-10	0-26	0-27	0-20	0-36	0-39	1
N	1921	912	684	2303	1313	1043	824	381	356	N
Mean	25.5	31.8	35.8	33.8	42.0	43.7	41.2	46.9	47.3	Mean
SD	8.9	8.6	7.9	8.9	5.4	4.9	6.3	2.9	2.7	SD

[a] Data derived from Form A standardization sample. Since total scores on Forms A and B were found to be equivalent, the norms presented here may be used for both Form A and Form B.

results of testing with both forms are explored in the Interpretation section of this Manual.

As an aid in interpreting local test results, the per-cent-passing figures for a particular classroom (or group of classrooms) may be compared with percentages given in the appropriate one of the four Tables, 5 through 8. The percentage of students in a classroom group who pass a given item on one of the forms may be compared with the percentage of students in the standardization sample who passed the same item, at the same grade, socioeconomic level, and time of administration.

Also of interest are the percentile equivalents of *BTBC* total raw scores, presented in Tables 9 and 10, for beginning-of-year and midyear testing, respectively. These data are based on the same groups on which the per-cent-passing figures for individual items on Form A (Tables 5 and 6) were computed. Since total scores on Forms A and B were found to be equivalent, the norms presented in Tables 9 and 10 may be employed for total scores obtained on either Form A or Form B.

The procedure for obtaining a child's percentile is to choose the norm table for the time of testing, locate the child's raw score in the appropriate column for grade and socioeconomic level, and read the percentile equivalent at the far right or left of the table. Each percentile point given in the tables represents a band of which the indicated percentile is approximately the midpoint. Thus, using the beginning-of-year norms, a child in kindergarten at a middle-socioeconomic-level school who obtains a raw score of 30 on the *BTBC* has a percentile rank of 40. This represents a band from 38 to 42, and indicates that his score surpasses at least 37 per cent of his group, and is surpassed by at least 58 per cent.[a]

Means and standard deviations based on the scores obtained by the Form A standardization sample are shown beneath the percentiles. A comparison of Tables 9 and 10 reveals that the middle- and high-socioeco-

[a] The percentile designations are the midpoints of bands which are five percentile units wide. The zones differ somewhat at the extremes. Thus, a percentile of 5 includes 4 through 7; 3 includes 2 and 3; 1 stands for the first percentile only. Similarly, the 95th percentile includes 93 through 96; 97 includes 97 and 98; 99 stands for only the 99th percentile.

Table 10. Percentile Equivalents of Raw Scores, by Grade and Socioeconomic Level[a]
(N=2647 Children Tested at Midyear)

Percentile	Kindergarten, Socioeconomic Level: Low	Middle	High	Grade 1, Socioeconomic Level: Low	Middle	High	Grade 2, Socioeconomic Level: Low	Middle	High	Percentile
99	45-50	48-50	50	49-50	50	—	50	—	—	99
97	43-44	47	49	48	—	50	49	50	—	97
95	40-42	46	48	47	49	—	—	—	50	95
90	37-39	44-45	47	46	48	49	48	—	—	90
85	36	43	46	45	—	—	—	—	—	85
80	35	42	45	44	47	—	47	49	—	80
75	34	41	44	43	—	48	—	—	—	75
70	32-33	40	43	42	—	—	46	—	49	70
65	31	39	42	—	46	—	—	48	—	65
60	30	38	—	41	—	47	45	—	—	60
55	29	37	41	—	45	—	—	—	—	55
50	28	36	40	40	—	—	44	47	—	50
45	27	35	39	39	44	46	—	—	48	45
40	26	34	—	38	—	—	43	—	—	40
35	25	33	38	37	43	45	—	46	—	35
30	24	31-32	36-37	36	—	—	42	—	47	30
25	23	30	35	35	42	44	41	45	—	25
20	21-22	28-29	34	34	41	43	40	—	46	20
15	19-20	26-27	32-33	32-33	39-40	42	39	44	—	15
10	15-18	24-25	29-31	31	37-38	41	36-38	42-43	45	10
5	13-14	19-23	25-28	29-30	34-36	38-40	34-35	41	44	5
3	10-12	14-18	22-24	27-28	31-33	35-37	27-33	39-40	41-43	3
1	0-9	0-13	0-21	0-26	0-30	0-34	0-26	0-38	0-40	1
N	162	453	250	276	413[b]	280	222	349	242	N
Mean	28.4	35.3	39.4	39.2	43.8	45.6	43.5	46.7	47.8	Mean
SD	8.1	8.0	6.5	5.5	4.5	3.7	5.0	2.7	2.6	SD

[a] Data derived from Form A standardization sample. Since total scores for Forms A and B were found to be equivalent, the norms presented here may be used for both Form A and Form B.

[b] One first-grade classroom group was excluded from the percentile computations. See footnote a, Table 6.

nomic-level grade 1 samples, and all of the grade 2 samples, obtained high mean scores with relatively little spread of scores. Even at the beginning of the school year, the majority of pupils in these groups will have mastered most of the basic concepts included on the test. Thus, for many groups of primary-grade pupils, the *BTBC* will be an easy and enjoyable experience. The test may, however, be quite useful for high-scoring groups. Its primary use at these levels lies in identifying the occasional concepts not well known to a particular group and the occasional pupils not well equipped to keep up with their peers.

It is interesting to note that the greatest increases in average scores from beginning of year to midyear tend to occur in kindergarten and grade 1, and at the lower socioeconomic levels. The groups that score lowest on the test in the fall (and therefore have the most to learn) tend to make the largest gains in mastery of concepts over the course of approximately a half-year.

In summary, Tables 5 through 10 should aid teachers in evaluating the performance of children on the *BTBC*. The teacher should bear in mind the fact that although the normative samples reflect a wide geographic distribution, they were not selected to be closely representative of United States school children as a whole. However, if the *BTBC* is used primarily to identify *children* in need of remedial concept training, or *concepts* unfamiliar to a large proportion of children, these uses do not require comparison with normative statistics.

Reliability

The reliability of a test refers to its consistency of measurement—that is, the accuracy of scores on the test, or the extent to which the scores are free of chance error. Reliability is normally expressed in either of two ways—as a reliability coefficient, or as a standard error of measurement. Split-half reliability coefficients and standard errors of measurement were both computed for Form A for the midyear standardization sample, and are presented in Table 11. Corresponding reliability statistics were computed for Form B, based on the scores of the pupils participating in the first testing session of the equivalence study who took Form B first; the resulting coefficients are also reported in Table 11.

The reliability coefficient may range from zero through 1.00, with higher values indicating greater reliability. The split-half reliability coefficients for the total score on Form A range from .68 to .90, while the corresponding coefficients for the total score on Form B range from .12 to .94. The coefficient of .12 is, of course, quite low. It was obtained for the grade 2, high-socioeconomic-level sample, a group which had a mean total score of 48.5 and a standard deviation of 0.9. Such a group forms the extreme upper end of the ability groups for which the *BTBC* has been presumed applicable. At this level, the value of the *BTBC* would seem to lie only in the identification of children who are far below the group's average ability.

One limitation of the reliability coefficient is its dependence, in part, on the variability of the test scores from which it was computed. The standard error of measurement (SE_m) provides an estimate of the amount of error associated with a test score. The SE_m is used to define a band around the "true" score within which about two-thirds of the raw scores would fall if a pupil were tested a very large number of times (neglecting the effects of practice, fatigue, and the like). For example, the Form A SE_m of 2.5 for the kindergarten children from high-socioeconomic-level schools indicates that the chances are about two out of three that the score obtained by an individual child from such a group is within 2.5 raw-score points of his true score. The smaller the standard error of measurement, the more confidence one may have in the accuracy of the test score.

The standard errors of measurement obtained for *BTBC* total scores appear to be essentially comparable for Form A and Form B.

Alternate form reliability is reflected in coefficients of correlation between two forms of a test, and indicates the level of consistency of measurement across these two forms. These coefficients are reported in Table 12 for total scores on Forms A and B. The sample consists

Table 11. Split-Half Reliability Coefficients and Standard Errors of Measurement, by Form

	Kindergarten			Grade 1			Grade 2		
	Socioeconomic Level:			Socioeconomic Level:			Socioeconomic Level:		
	Low	Middle	High	Low	Middle	High	Low	Middle	High
Form A[a]									
N	162	453	250	276	413	280	222	349	242
Mean	28.4	35.3	39.4	39.2	43.8	45.6	43.5	46.7	47.8
SD	8.1	8.0	6.5	5.5	4.5	3.7	5.0	2.7	2.6
r_{11}[b]	.86	.90	.85	.82	.82	.76	.82	.68	.73
SE_m[c]	3.0	2.5	2.5	2.3	1.9	1.8	2.1	1.5	1.4
Form B[d]									
N	39	72	47	60	88	70	65	76	53
Mean	31.6	40.5	36.6	39.8	45.7	44.7	44.9	46.2	48.5
SD	8.4	5.7	7.0	7.1	2.9	3.4	6.9	3.4	0.9
r_{11}[b]	.83	.78	.80	.88	.57	.60	.94	.62	.12
SE_m[c]	3.4	2.7	3.2	2.4	1.9	2.2	1.7	2.1	0.9

[a] Form A reliability sample consists of the 2647 children whose scores on Form A were used in deriving midyear norms (see Table 10).

[b] Split-half reliability coefficients, based on the correlation of scores on odd-numbered items with scores on even-numbered items, corrected by the Spearman-Brown formula. Similar reliability coefficients were computed for the total group of pupils at each of the three grades. For Form A, these coefficients were .90, .85, and .81 for the kindergarten, grade 1, and grade 2 samples, respectively. For Form B, the coefficients were .84, .83, and .87 for the kindergarten, grade 1, and grade 2 samples, respectively.

[c] Standard error of measurement computed by the formula, $SE_m = SD\sqrt{1 - r_{11}}$. For the total groups of pupils in the three grades—kindergarten, grade 1, and grade 2—the SE_m for Form A was 2.7, 2.1, and 1.7, respectively, and the SE_m for Form B was 3.0, 2.2, and 1.7, respectively.

[d] Form B reliability sample consists of all children in the equivalence study who were administered Form B at the first testing session. This group is approximately one-half of the sample described in Table 4.

of those students in the equivalence study who were present for both the first and the second testing session. The interval between testings was always more than one day but less than one week. Alternate form coefficients tend to be slightly lower in magnitude than those obtained by the split-half method; these ranged from .55 to .92, with a median of .76.

Validity

An essential aspect of the validity of a test is how well the test performs its intended work. For the *Boehm Test of Basic Concepts*, like any other test of educational achievement or mastery, validity is primarily a matter of the relevance of the test content to the school curriculum. This type of validity is usually called *content* validity. (As an example, a test of typewriting speed and accuracy would have high content validity if used at the start of an intermediate typewriting course.) In the case of the *BTBC*, the test items were selected from relevant curriculum materials and represent concepts basic to understanding directions and other oral communications from teachers at the preschool and primary-grade level.

Table 12. Alternate Form Reliability Coefficients[a]

	Kindergarten			Grade 1			Grade 2		
	Socioeconomic Level:			Socioeconomic Level:			Socioeconomic Level:		
	Low	Middle	High	Low	Middle	High	Low	Middle	High
N	76	134	81	123	163	137	120	150	151[b]
r_{AB}[c]	.58	.55	.78	.90	.68	.76	.92	.78	.63
Form A: Mean	31.1	39.7	37.2	40.0	45.7	45.5	44.3	47.3	47.6
SD	9.6	5.8	5.8	7.1	3.4	3.8	6.4	2.6	1.9
Form B: Mean	34.2	40.7	37.5	40.2	45.9	45.3	44.4	47.3	47.9
SD	7.4	5.2	6.8	6.6	3.1	3.5	6.3	3.0	1.8

[a] Coefficients based on all children participating in the equivalence study who were administered both forms of the test. (The N's are lower than those reported in Table 4 because Table 4 includes some children who were only present for one testing session.) For each form separately, data from the first and second testings were combined.

[b] Through an administrative error, one group that should have been tested with Form B first was given Form A first. However, scores obtained by this group were included in calculating the alternate form reliability coefficients. See Table 4, footnote b.

[c] Alternate form coefficients were computed for the total group of pupils at each of the three grade levels. These coefficients were .72, .87, and .88 for the kindergarten, grade 1, and grade 2 samples, respectively.

APPENDIX

Table for Determining Per Cent Passing an Item

Number of Children Answering Item Correctly	Number of Children Tested: 11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	Number of Children Answering Item Correctly
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	9	8	8	7	7	6	6	6	5	5	5	5	4	4	4	4	4	4	3	3	3	3	3	3	3	3	3	3	3	2	1
2	18	17	15	14	13	12	12	11	11	10	10	9	9	8	8	8	7	7	7	7	6	6	6	6	6	6	5	5	5	5	2
3	27	25	23	21	20	19	18	17	16	15	14	14	13	12	12	12	11	11	10	10	10	9	9	9	9	8	8	8	8	8	3
4	36	33	31	29	27	25	24	22	21	20	19	18	17	17	16	15	15	14	14	13	13	12	12	12	11	11	11	11	10	10	4
5	45	42	38	36	33	31	29	28	26	25	24	23	22	21	20	19	19	18	17	17	16	16	15	15	14	14	14	13	13	12	5
6	55	50	46	43	40	38	35	33	32	30	29	27	26	25	24	23	22	21	21	20	19	19	18	18	17	17	16	16	15	15	6
7	64	58	54	50	47	44	41	39	37	35	33	32	30	29	28	27	26	25	24	23	23	22	21	21	20	19	19	18	18	18	7
8	73	67	62	57	53	50	47	44	42	40	38	36	35	33	32	31	30	29	28	27	26	25	24	24	23	22	22	21	21	20	8
9	82	75	69	64	60	56	53	50	47	45	43	41	39	38	36	35	33	32	31	30	29	28	27	26	26	25	24	24	23	22	9
10	91	83	77	71	67	62	59	56	53	50	48	45	43	42	40	38	37	36	34	33	32	31	30	29	29	28	27	26	26	25	10
11	100	92	85	79	73	69	65	61	58	55	52	50	48	46	44	42	41	39	38	37	35	34	33	32	31	31	30	29	28	28	11
12		100	92	86	80	75	71	67	63	60	57	55	52	50	48	46	44	43	41	40	39	38	36	35	34	33	32	32	31	30	12
13			100	93	87	81	76	72	68	65	62	59	57	54	52	50	48	46	45	43	42	41	39	38	37	36	35	34	33	32	13
14				100	93	88	82	78	74	70	67	64	61	58	56	54	52	50	48	47	45	44	42	41	40	39	38	37	36	35	14
15					100	94	88	83	79	75	71	68	65	62	60	58	56	54	52	50	48	47	45	44	43	42	41	39	38	38	15
16						100	94	89	84	80	76	73	70	67	64	62	59	57	55	53	52	50	48	47	46	44	43	42	41	40	16
17							100	94	89	85	81	77	74	71	68	65	63	61	59	57	55	53	52	50	49	47	46	45	44	42	17
18								100	95	90	86	82	78	75	72	69	67	64	62	60	58	56	55	53	51	50	49	47	46	45	18
19									100	95	90	86	83	79	76	73	70	68	66	63	61	59	58	56	54	53	51	50	49	48	19
20										100	95	91	87	83	80	77	74	71	69	67	65	62	61	59	57	56	54	53	51	50	20
21											100	95	91	88	84	81	78	75	72	70	68	66	64	62	60	58	57	55	54	52	21
22												100	96	92	88	85	81	79	76	73	71	69	67	65	63	61	59	58	56	55	22
23													100	96	92	88	85	82	79	77	74	72	70	68	66	64	62	61	59	58	23
24														100	96	92	89	86	83	80	77	75	73	71	69	67	65	63	62	60	24
25															100	96	93	89	86	83	81	78	76	74	71	69	68	66	64	62	25
26																100	96	93	90	87	84	81	79	76	74	72	70	68	67	65	26
27																	100	96	93	90	87	84	82	79	77	75	73	71	69	68	27
28																		100	97	93	90	88	85	82	80	78	76	74	72	70	28
29																			100	97	94	91	88	85	83	81	78	76	74	72	29
30																				100	97	94	91	88	86	83	81	79	77	75	30
31																					100	97	94	91	89	86	84	82	79	78	31
32																						100	97	94	91	89	86	84	82	80	32
33																							100	97	94	92	89	87	85	82	33
34																								100	97	94	92	89	87	85	34
35																									100	97	95	92	90	88	35
36																										100	97	95	92	90	36
37																											100	97	95	92	37
38																												100	97	95	38
39																													100	98	39
40																														100	40

DIRECTIONS: Locate the appropriate column for the number of children tested. Locate the number of children answering the item correctly (at the far left or right of the table). Move across this row to the appropriate column to find the per cent passing the item. For example, if 17 out of a class of 22 children answer a certain item correctly, the per cent passing the item is read as 77.

THE PSYCHOLOGICAL CORPORATION
INCORPORATED IN 1921 ®

Appendix 5

The Boehm Test of Basic Concepts Class Record Form.

THE PSYCHOLOGICAL CORPORATION ®
INCORPORATED IN 1921

Appendix 5

The Boehm Test of Basic Concepts Class Record Form.

FORM A

BOEHM TEST of basic concepts

Ann E. Boehm

CLASS RECORD FORM

SCHOOL ______________________ CLASS __________

TEACHER ______________________

DATE OF TESTING ______________________

Printed in U. S. A.
The Psychological Corporation, 304 East 45th Street, New York, N. Y. 10017
71-2375

DIRECTIONS: Enter the names of the children tested. In the column below a child's name, record his test performance by entering a check mark (✓) for each item answered correctly. Do this for each child. Then count the ✓'s in each row and column, and enter the appropriate totals in the shaded row and shaded column. Enter the total of the shaded row in the space marked SUM. The total of the shaded column should be the same number.

To compute the entries for the column marked PER CENT PASSING, for each item, divide the total number of children answering it correctly by the total number of children tested. (The Appendix can be used to determine PER CENT PASSING for classes of 11 to 40 children.) To compute the CLASS AVERAGE, divide the SUM by the total number of children tested. See the Manual for determining PERCENTILES. (For further scoring and recording directions, see reverse side of this form.)

CONTEXT CATEGORIES

S = Space (location, direction, orientation, dimensions)
Q = Quantity (and number)
T = Time
M = Miscellaneous

CHILD'S NAME

CONCEPT	KEY	CONTEXT CATEGORY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	TOTAL NUMBER OF CHILDREN ANSWERING CORRECTLY	PER CENT PASSING % = Number answering correctly / Number tested
BOOKLET 1 1. Top		S																																
2. Through		S																																
3. Away from		S																																
4. Next to		S																																
5. Inside		S																																
6. Some, not many		Q																																
7. Middle		S																																
8. Few		Q																																
9. Farthest		S																																
10. Around		S																																
11. Over		S																																
14. Between		S																																
15. Whole		Q																																
16. Nearest		S																																
17. Second		Q																																
18. Corner		S																																
19. Several		Q																																
20. Behind		S																																
21. Row		S																																
22. Different		M																																
23. After		T																																
24. Almost		Q																																
25. Half		Q																																
BOOKLET 2			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
26. Center		S																																
27. As many		Q																																
28. Side		S																																
29. Beginning		T																																
30. Other		M																																
31. Alike		M																																
32. Not first or last		Q																																
33. Never		T																																

Item		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Sum*	Class Average†
22. Different	M																																
23. After	T																																
24. Almost	Q																																
25. Half	Q																																
BOOKLET 2		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
26. Center	S																																
27. As many	Q																																
28. Side	S																																
29. Beginning	T																																
30. Other	M																																
31. Alike	M																																
32. Not first or last	Q																																
33. Never	T																																
34. Below	S																																
35. Matches	M																																
36. Always	T																																
37. Medium-sized	Q																																
38. Right	S																																
39. Forward	S																																
40. Zero	Q																																
41. Above	S																																
42. Every	Q																																
43. Separated	S																																
44. Left	S																																
45. Pair	Q																																
46. Skip	M																																
47. Equal	Q																																
48. In order	S																																
49. Third	Q																																
50. Least	Q																																
SCORE (Total Number of Items Answered Correctly)		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Sum*	Class Average†
PERCENTILE																																	
NORMS USED ______																																	

TOTAL NUMBER OF CHILDREN TESTED

N=

* Sum of shaded row should equal sum of shaded column. If not, check work for error.

† Class average = $\frac{\text{Sum}}{\text{Number tested}}$

How to Score and Record the Results

The Class Record Form serves as both a scoring key and an interpretive aid to the teacher. It is designed to speed recording and facilitate analysis of a class's test results. If there are more than 30 children in the class, two Class Record Forms should be fastened together so that the entries in the shaded column headed "Total Number of Children Answering Correctly" and the adjacent column headed "Per Cent Passing" will be based on the scores of the entire class.

Down the left-hand side of the form the concepts covered by the test are listed by item number, in the sequence in which they appear in the test booklets. Each concept is accompanied by a miniature reproduction of the set of pictures by which it is tested, appropriately marked to show the correct response. The Class Record Forms for both forms of the test are alike with respect to the concepts listed and the sequence of listing but of course they have different picture reproductions, each marked with the correct response. An illustration of the Form A Class Record Form is provided on page 13.

Across the top of the form are spaces for entering the names of the children tested. The form has been marked off into boxes for recording the response of each child to each item. If the same children are given both forms of the test (for example, Form A at the beginning and Form B at the end of the school year), it should prove helpful in comparing the results of the two testings if the names of the children have been entered in the same sequence on both Class Record Forms.

Two suggested procedures for using the Class Record Form are described below. With either plan it will require approximately one hour to complete the form for a group of 30 children. About 20 additional minutes are needed to cross-check the column and row sums and to enter the per cent passing for each of the 50 items.

Plan A

The following procedure is recommended for scoring Booklets 1 and 2 together:

1. Place together the two booklets for each child (with Booklet 2 following Booklet 1). Then stack the pairs of booklets in any order desired (for example, alphabetically).

2. Write the children's names, in the order in which the booklets are stacked, in the spaces provided for this purpose at the top of the Class Record Form.

3. Starting with the pair of booklets for the first child, go through them item by item, making a check mark (✓) in the appropriate space of that child's column for each item answered *correctly*. Use the miniature scoring key on the Class Record Form for this purpose.[4] Make no check marks at all for items answered incorrectly or ambiguously, or for items that have been omitted. Repeat this procedure for each child tested.

4. Count the check marks in each *row* and record the total in the *shaded column* marked "Total Number of Children Answering Correctly" on the right-hand side of the Class Record Form.

5. Count the check marks in each *column* and record the total in the *shaded row* labeled "Score (Total Number of Items Answered Correctly)" at the bottom of the form.

6. Add the numbers in the *shaded row*—"Score (Total Number of Items Answered Correctly)"—and record this figure in the box labeled "Sum." It should agree with the total of the *shaded column* ("Total Number of Children Answering Correctly").[5]

7. Divide the "Sum" by the total number of children tested, and record the result in the box labeled "Class Average." This quotient will represent the group's average score on the total test.

8. To determine the per cent of children in the class who answered each item correctly, divide each entry in the shaded column labeled "Total Number of Children Answering Correctly" by the total number of children tested. Record these per cents in the column at the right headed "Per Cent Passing." For classes of 11-40 pupils, the table in the Appendix may be used to determine conveniently the per cent of children in the class who answered an item correctly. Directions are printed on the table.

9. A row for entering percentiles is provided at the bottom of the Class Record Form. These may be obtained from the Norms section of this Manual, or from norms based on local data.

Plan B

If the teacher wishes to score Booklets 1 and 2 *separately*, the following is suggested as an alternative procedure:

1. Arrange all copies of Booklet 1 in the order desired. Then arrange all copies of Booklet 2 in this same order.

2. Write the children's names, in the order in which the Booklet 1's are stacked, in the spaces provided for this purpose at the top of the Class Record Form.

3. Fold the Class Record Form back along the line between items 25 and 26, leaving only the upper part, containing items 1-25, exposed. For each child, enter check marks for correct answers to items 1-25 (Booklet 1) in the manner previously described.

4. To score Booklet 2, fold the Class Record Form so that the fold already made along the line between items 25 and 26 is placed along the line immediately above item 1, leaving only the part containing items 26-50 exposed. Enter check marks for correct answers to items 26-50 (Booklet 2). *Be careful to use the correct column for each child.* If desired, the number appearing at the bottom of each name space on the Class Record Form can be entered on the cover of the two booklets marked by that child.

5. Complete the Class Record Form by following steps 4 through 9 of Plan A.

When the Class Record Form has been completed, the examiner may wish to fold the area containing the last two rows ("Score" and "Percentile") under, and then align this fold with the bottom of the "Child's Name" section at the top of the record form. This will facilitate the reading of each child's score.

[4] Alternatively, the teacher may mark a set of the *BTBC* booklets with the correct answers, and use them in scoring the test.

[5] If not the same, check for counting errors.

Appendix 6

Adapted Directions for Form A

Booklet 1

1. "Look[1] at these pieces of paper with stars on. Find[2] the paper with the star at the top... Point[2] to the paper with the star at the top.

2. "Look at the beads and strings. Find the bead that has a string through it... Point to the bead with a string through it.

"Look at the table and boxes. Find the box that is away from the table... Point to the box that is away from the table.

4. "Look at the toys. Find the toy that is next to the truck, (the lorry.[3])... Point to the toy that is next to the truck/lorry.

5. "Look at these pictures of the house and boy". (Indicate item 5 but avoid pointing at any particular picture.) "Find the house with the boy inside it... Point to the house with the boy inside it.

6. "Look at the boxes and marbles. Find the box that has some but not many marbles... Point to the box that has some but not many marbles.

1 In each case, the interrogative phrase "See the...?" may be substituted.

2 In each case, "Show me..." or "Touch the..." may be substituted.

3 Add "the lorry" for 3 and 4 year olds.

7. "Look at the flowers. Find the flower that is in the <u>middle</u>... Point to the flower that is in the <u>middle</u>.

8. "Look at the plates of cakes. Find the plate with a <u>few</u> cakes... Point to the plate with a <u>few</u> cakes.

9. "Look at the boats. Here's the shore". (Point to it.) "Find the boat that is <u>farthest</u> from the shore... Point to the boat that is <u>farthest</u> from the shore.

10. "Look at the boxes and circles" (For younger children substitute the word 'balls'.) Find the box that has circles (balls) <u>around</u> it... Point to the box with circles <u>around</u> it.

11. "Look at the balloons and the tree. Find the balloon that is <u>over</u> the tree... [1]Point to the balloon that is <u>over</u> the tree.

12. "Look at the doors. Find the door that is <u>widest</u>. ...the <u>widest</u> door.

13. "Look at the boxes of eggs. Find the box with the <u>most</u> eggs. ...the box with the <u>most</u> eggs.

14. "Look at the jars, cups and spoons. Find the thing that is <u>between</u> the spoons. ...the thing <u>between</u> the spoons.

[1] As the test progresses it will become obvious when the repetition of "point to" or "show me" has become superfluous. The repetition of the concept word(s) should be maintained except where the child has already pointed to the correct answer.

15. "Look at the cakes. Find the cake that is whole, the whole cake. ...the whole cake.

16. "Look at the boys going to school. Find the boy who is nearest the door. ...the boy nearest the door.

17. "Look at the animals walking in a line. Find the second animal. ...the second animal.

18. "Look at the glasses on the table. Find the glass at the corner of the table. ...the glass at the corner.

19. "Look at the pictures of animals. Find the picture with several rabbits. ...the picture with several rabbits.

20. "Look at the settee and the toys. Find the toy that is behind the settee. ...the toy behind the settee.

21. "Look at the trees. Find the group where all the trees are in a row. ...the trees all in a row.

22. "Look at the bricks. Find the pile of bricks that is different from the others. ...the bricks that are different.

23. "Look at the pictures of a girl. See, here she's having her hair cut. Show me the girl after her hair was cut. ...the girl after her hair was cut.

24. "Look at the bottles. Find the bottle that is almost empty. ...the bottle that is almost empty.

25. "Look at the pies. Find the pie that is half gone. ...the pie that's half gone."

Booklet 2

26. "Look at the circle (the ring) and boxes. Find the box at the centre of the circle (ring). ...the box at the centre of the circle (ring).

27. "Look at the box of marbles and the groups of marbles. Find the group (set) that has as many marbles as the box has. ...the group (set) with as many marbles as the box has.

28. "Look at the box and the circles (balls). Find the circle (ball) at the side of the box. ...the circle at the side of the box.

29. "Look at the trees and squirrels. Find the squirrel that is beginning to climb a tree. ...the squirrel beginning to climb.

30. "Look at the puddings. One is an ice-cream and one is a piece of pie. Find the other one. ...the other pudding.

31. "Look at the shapes. Find the shapes that are alike. ...the shapes that are alike.

32. "Look at the cars going into the tunnel. Find the car that is not the first or the last. ...the car that's not the first or the last.

33. "Look at the chair, the apple and the biscuits. Show me what a child (boy/girl) should never eat. ...what a child should never eat.

34. "Look at the table. Put your finger below the table. ...below the table.

35. "Look at the boxes and balls. Find the ball that matches one of the boxes. ...the ball that matches one of the boxes.

36. "Look at the dog, the book and the ear. Show me the one a child always has. Which does a child always have?

37. "Look at the fish. Find the fish that is middle-sized[1]. ...[1]the middle-sized fish.

38. "Look at the boxes and the line. Find the box at the right end of the line. ...the box on the right.

39. "Look at the boys. Find the boy who is bending forward. ...the boy bending forward.

40. "Look at the boxes of sweets. Find the box without any sweets. ...the box without any sweets.

41. "Look at the cloud and aeroplanes. Find the aeroplane above the cloud. ...the aeroplane above the cloud.

42. "Look at the pictures of bowls and spoons. Find the picture that has a spoon in every bowl. ...the picture a spoon in every bowl.

[1] For seven year olds: "... middle-sized, the medium-sized fish.

43. "Look at the beads. Find the beads that are separated. ...the beads that are separated.

44. "Look at the birds. Find the bird on the left. ...the bird on the left.

45. "Look at the candles. Find a pair of candles. ...a pair of candles.

46. "Look at the boxes. Put your finger on the box with the X. Now miss[1] a box and touch the next box. ...miss[1] a box and touch the next box.

47. "Look at the pictures of lollipops. Find two pictures with equal numbers of lollipops. ...two pictures with equal numbers.

48. "Look at the boxes of circles. Find the box where the circles are in order from large to small. ...the box where the circles are in order from large to small.

49. "Look at the teacher and the children. Find the third child from the teacher. ...the third child from the teacher.

50. "Look at the stars. Find the group that has the least stars. ...the group with the least stars.

[1] A better modification would be miss out. (See p.335)

Appendix 7

Adapted Directions for Form B

See footnotes to Appendix 6.

Booklet 1

1. See Appendices 4 and 6.

2. See Appendices 4 and 6.

3. "Look at the baby and the bricks. Find the brick that is away from the baby."

4. "Look at these animals". (Indicate item 4 but avoid pointing to any particular animal.) "Find the animal that is next to the rabbit... Point to the animal next to the rabbit.

5. "Look at the boxes and balls. Find the box with balls inside it... Point to the box with balls inside it.

6. "Look at the vases of flowers. Find the vase that has some but not many flowers... Point to the vase that has some but not many flowers.

7-13. See Appendices 4 and 6 (i.e. use the manual directions for Form B but follow the directions for individual testing which are adopted with Form A).

14. "Look at these pictures of toys. Find the picture that has a bear between two bricks. ... a bear between two bricks.

Appendix 7

Adapted Directions for Form B

See footnotes to Appendix 6.

Booklet 1

1. See Appendices 4 and 6.

2. See Appendices 4 and 6.

3. "Look at the baby and the bricks. Find the brick that is away from the baby."

4. "Look at these animals". (Indicate item 4 but avoid pointing to any particular animal.) "Find the animal that is next to the rabbit... Point to the animal next to the rabbit.

5. "Look at the boxes and balls. Find the box with balls inside it... Point to the box with balls inside it.

6. "Look at the vases of flowers. Find the vase that has some but not many flowers... Point to the vase that has some but not many flowers.

7-13. See Appendices 4 and 6 (i.e. use the manual directions for Form B but follow the directions for individual testing which are adopted with Form A).

14. "Look at these pictures of toys. Find the picture that has a bear between two bricks. ... a bear between two bricks.

15. "Look at the apples. Find the apple that is whole, the whole apple. ...the whole apple."

16. See Appendices 4 and 6.

17. "Look at these trucks, lorries, and the stop sign. Find the second truck from the sign. ...the second truck from the sign."

18. See Appendices 4 and 6.

19. "Look at the knives, forks and spoons. Find the picture that has several spoons. ...the picture with several spoons.

20. "Look at the boys and the trailer. Find the boy who is behind the trailer. ...the boy behind the trailer.

21. "Look at the pictures of bottles. Find the bottles which are all in a row. ...the bottles all in a row."

22. See Appendices 4 and 6.

23. "Look at the pictures of a piece of wood. Find the piece of wood after it was cut. ...the wood after it was cut.

24. See Appendices 4 and 6.

25. See Appendices 4 and 6.

Booklet 2

26-32 See Appendices 4 and 6.

33. "Look at the lamp, the watch and the shoe. Find the thing that a child never wears. ...the thing a child never wears.

34. "Look at the seat and the birds. Find the bird below the seat. ...the bird below the seat.

35. "Look at the T-shirts and shorts. Find the shorts that match one of the shirts. ...the shorts that match one of the shirts."

36. See Appendices 4 and 6.

37. "Look at the butterflies. Find the butterfly that is middle-sized[1]. ...the butterfly that is middle-sized.

38. See Appendices 4 and 6.

39. "Look at the chicks. Find the chick that is bending forward. ...the chick bending forward.

40. "Look at the rabbits and carrots. Find the rabbit without any carrots. ...the rabbit without any carrots."

41-44 See Appendices 4 and 6.

1 See footnote to item 37, Appendix 6.

45. "Look at the dolls. Find a pair of dolls. ...a pair of dolls."

46. "Look at the circles (rings). Put your finger on the ring (circle) with the X. Miss[1] a ring and touch the next ring. ...Miss[1] a ring and touch the next one.

47. "Look at the stars. Find two pictures that have equal numbers of stars. ...two pictures with equal numbers of stars.

48. See Appendices 4 and 6.

49. "Look at the shop and the houses. Find the third house from the shop. ...the third house from the shop.

50. "Look at the ice-creams. Find the picture with the least ice-creams. ...the picture with the least ice-creams.

[1] A better modification would be miss out. (See p.355)

Appendix 8

Parental Occupations and SES Groupings[1]

Below are examples of the occupations of parents of the children in the sample population. These are listed under the socio-economic group to which they were assigned for this research. Grades in parenthesis are the Registrar-General's classification; where more than one grade is given, insufficient information was available to make a precise classification.

SES Group 1: a sample of the occupations

Area sales manager (II)

Biochemist (I or II)

Builder, own business (II)

Chartered accountant (I)

Chiropodist (II)

Club secretary (II)

Coach proprietor (II)

Company director (I)

Customs officer (II)

Data processor (II)

Disabled Resettlement Officer (II)

District supervisor, M. of Defence (II)

Doctor (G.P.) (I)

Farmer (II)

Farm manager (II)

Fisheries inspector, Y.W.A. (II)

[1] See Chapter 8, i, 2.14

Garage owner (II)

Geologist, N.C.B. (II)

Hydraulics test engineer (II)

Landscape gardener, own firm (II)

Manager, B.S.C. (I or II)

Managing director, finance company (I)

Marine engineer (II)

Medical rep., graduate (II)

Metallurgist, B.S.C. (I)

Minister (clergy) (I)

N.C.B. contracts officer (II)

N.C.B. deputy (II)

N.C.B. manager (II)

Nurse (II)

Pharmacist (II)

Planner (I or II)

Publican

Quarry manager (II)

Restaurant owner (II)

River pilot (II)

R.A.F. Officer (I or II)

Senior executive at County Hall (I or II)

Solicitor (I)

Teacher (II)

University Lecturer (I)

SES Group 2: a sample of the occupations

All the occupations in this group are classified by the Registrar-General as either III skilled manual or III white collar.

Agricultural representative

Assistant manager in shop

Bank clerk

Blacksmith

Blast furnaceman

Bricklayer

Bus driver

Cable jointer

Caravan fitter

Carpenter

Central heating engineer

Clerk in Post Office

Contract tunneller

Crane driver

Draughtsman

Driving instructor

Electrician

Fitter

Fireman

Garage mechanic

Glass inspector

Head barman

Insurance agent

HGV driver

Joiner

Leather technician

Machine setter

Manager of off-licence

Miller

N.C.B. electrician

Pipe fitter

Pit winder

Plumber

Policeman

Quality Control, B.S.C.

R.A.F. aircrew

Secretary

Shunter, B.S.C.

Steel erector

Subsidence claims in Estate Agents

Supervisor, Nypro Chemicals

Telecom engineer

Tool maker

Tool setter

Turbine technician

Van driver

Warehouse stock controller

Welder

Woodcarver, self-employed

SES Group 3: a sample of the occupations

Baths attendant (IV)

Builder's labourer (V)

Barmaid (IV)

B.S.C. coke ovens (IV)

B.R. lines railman (IV)

Bus conductor (IV)

Deckhand on fishing trawler (IV)

Docker (V)

Farm labourer (IV)

Fettler (IV)

Groundsman (IV)

Hotel chambermaid (V)

Machine operator (IV)

Machinist (IV)

Navvy, Humber Bridge (V)

Packer (IV)

Paint sprayer (IV)

Postman (IV)

Profile burner (IV)

Process operator (IV)

Refuse collector (V)

Sorter at Rockware (IV)

Unemployed, long-term (V)

Wiredrawer (IV)

Yardsman (V)

SES Group 4: a sample of the occupations

Glass worker (III or IV)

Manual (III, IV or V)

Shipyard worker (III, IV or V)

Steelworker (III, IV or V)

Traffic department (III, IV or V)

Quarry worker (III, IV or V)

SES Group 5: a sample of the occupations

Accountant (I, II or III)

B.S.C. computer department (II or III)

Engineer (II or III)

Surveyor or draughtsman (II or III)

Appendix 9

Dear

It is now about   months since you kindly tolerated me testing some of the children in your school. It would help in validating the Boehm Test if you or your staff would be good enough to provide me with information on the children's present achievements preferably without referring to the results I gave you.

I should be grateful if the completed enclosures could be returned within the next few weeks. I accept that you may decide that this would over-burden hard-working teachers. In that case please return the forms uncompleted in the enclosed envelope.

I also enclose the names of children for whom nothing was known about their parents' occupations. If you now happen to have this information I should be grateful for it. Recent unemployment is not relevant: the father's or/and mother's trade or profession when employed is what I am seeking.

Yours sincerely,

Copy of the letter sent to the headteachers of schools which were used in the predictive validity study.

Appendix 10

Form used for Rating Spoken Language

SPOKEN LANGUAGE

1. Grade +, Av or minus.

   + : well above average for age group in this school
   Av : in the broad band of average group in this school
   minus : well below average for age group in this school

   If you wish to indicate that a child may be marked differently if seen against the national average please add +, Av or minus in the column headed 'Comments'.

2. The 'Comments' column may be left blank or used in any way you wish.

3. Language Please disregard speech impediments and pronunciation. Take into account the child's understanding of classroom instructions, ability to frame questions, ability to maintain a conversation with children and adults, adequacy of vocabulary and the ability to describe an event coherently. If a child is rated low on language and it is clear that this is because of an emotional problem, rather than lack of ability, please indicate this in the 'Comments' column.

Name	Date of birth	Grade	Comments

Appendix 11

Form used for Rating Mathematics

1. Grade +, Av or minus.

   + : well above average for age group in this school
   Av : in the broad band of average group in this school
   minus : well below average for age group in this school

   If you wish to indicate that a child may be marked differently if seen against the national average please add +, Av or minus in the column headed 'Comments'.

2. The 'Comments' column may be left blank or used in any way you wish.

3. MATHEMATICS

   A Computational skills Please take into account oral and written skills in counting and computation with and without concrete aids.

   B Concepts of space, quantity and time. Please consider the child's ability to understand and apply such concepts as behind, few, full, higher, most, smallest, wider etc. as well as practical work in measuring (linear, capacity and weight). Understanding of various time concepts such as after, next week, yesterday and the passage of time are of more relevance than clock recognition.

Name	Date of birth	Grade A. Comp.	Grade B. Concepts	Comments

Appendix 12

Form used for Rating Reading Attainment

1. Mark +, Av or minus.

   + : well above average for age group <u>in this school</u>
   Av : in the broad band of average group <u>in this school</u>
   minus : well below average for age group <u>in this school</u>

   If you wish to indicate that a child may be marked differently if seen against the national average please add +, Av or minus in the column headed 'Comments'.

2. The 'Comments' column may be left blank or used in any way you wish.

3. <u>READING</u> Please take into account the child's abilities in word recognition, reading with understanding and general reading fluency. Should this not apply to the youngest children please mark 'Pre-reader'.

Name	Date of birth	Grade	Comments

Appendix 13

Reading Books Used in Content Validity Study

Breakthrough to Literacy by Mackay, D., Thompson, B. and Schaub, P. (1970)

London: Longman/Schools Council

level	books used[1]
Yellow	11 books
Red	12 books
Blue Set A	4 books
Set B	4 books
Green Set A	4 books
Set B	4 books
Set C	4 books
Red Set F ("Jokers")	4 books

Dominoes by Glynn, Dorothy M. (1972; 1974; 1976)

Oliver and Boyd

level	books used
Red	Books 1 - 6
Blue	Books 1 - 6
Green	Books 1 - 6
Pink	Books 1 - 6
Yellow	Books 1 - 6

Ladybird Key Words Reading Scheme by Murray, W. (1970)

Loughborough: Wills and Hepworth Ltd.

Introductory Books 1a, 2a, 3a

Easy Readers 1 - 6

Books 4a - 9a

[1] The books available to the author. In the majority of cases these were all the books in the series in print.

<u>Language in Action</u> by Morris, Joyce M. (1974)

London: Macmillan Educational

level	books used
Pre-Literary	26 books
1. Red	12 books
2. Green	8 books
3. Blue	Sue's Blue Hoop; The Old Farm Cart.

<u>Reading 360</u>. The Ginn Reading Programme (1978; 1980)

London: Ginn and Co. Ltd.

level	books used
1	Books 1 - 3
2	Books 1 - 3
3	Books 1 - 3
4	Books 1 - 3
5	Books 1 - 5
6	Book 3, Book 5, Book 6, Book 7
7	Books 1 - 7
8	Books 1 - 7

<u>Sparks</u> by Fisher, R.M., Hynds, M., Johns, A.M. and McKenzie, M.G. (1972)

London: Blackie & Son Ltd.

level	books used
1. Red	11 books
2. Blue	12 books
3. Green	16 books
4. Yellow	8 books
5. Pink	4 books
6. Purple	8 books
Sparks Bookshelf:	Stop that Dragon; The Fireball

Through the Rainbow by Bradburne, E.S. (1966)

Huddersfield: Schofield & Sims Ltd.

level	books used
Red	Books 1 - 3
Orange	Books 1 - 3
Yellow	Books 1 - 3
Green	Books 1 - 3
Blue	Books 1 - 3
Indigo	Books 1 - 3
Violet	Books 1 - 3
Rainbow	Books 1 - 3
Silver	Books 1 - 3
Gold	Books 1 - 3

Appendix 14

Mathematics Books Used in Content Validity Study

Alpha Mathematics 1 Goddard, T.R. and Grattidge, A.W. (1969)
Huddersfield: Schofield & Sims Ltd.

Mathematics for Younger Children. Biggs, E. (1971)
London: Macmillan

Early Mathematical Experience Matthews, G. and Matthews, J. (1978)
London: Schools Council/Addison-Wesley Ltd.

E.M.E. General Guide.
Home Corner. The Family.
Outdoor Activities. The Environment.
Outdoors Environment.
Passage of Time. Rhymes and Stories.
Space and Shape. Comparisons.
Towards Number. Toys and Games.
Water. Raw Materials.

Mathematics for Schools. Second Edition. Howell, A., Walker, R. and Fletcher, H. (1979) London: Addison-Wesley Ltd.

Level I : Books 1 - 7
Level II: Book 0, Book 1.

Maths Adventure 1. Pupils' Book. Stansfield, J. (1971)
London: Evans Bros. Ltd.

Nuffield Mathematics Project. The Nuffield Foundation. (1967; 1970)
London: W & R Chambers and John Murray.

Beginnings 1.
Pictorial Representation.
I Do and I Understand.
Mathematics Begins.
Computation and Structure 2.
Shape and Size 2.
Checking Up 1.

Peak Mathematics. Brighouse, A., Godbar, D. and Patilla, P. (1981)
Walton-on-Thames: Thomas Nelson & Son Ltd.

Infant Handbooks 1 and 2.
Book 0.
Books 1 and 2.

Primary Mathematics. A development through activity Scottish Primary Mathematics Group. (1975) London: Heinemann Educational Books Ltd.

Teachers' Notes. Stages 1 and 2.
Stage 2: Workbook 2.

Appendix 15

A random sample of 45 children (5 at each age level).

A record card was filed on each of the 1928 children in the population sample. (See Chapter 10). The information below is extracted from these cards.

Explanatory notes

1. (863) is the child's identification number.

2. 1st half-year/2nd half-year refer to the time of testing. (See Chapter 8, iii)

3. SES: Roman numericals indicate the Registrar-General's classification of occupation; Arabic numerals indicates the SES group used in this research. (See Chapter 8, i, 2.14) SPF indicates a single parent.

4. School: F indicates a First School.
   Inf. indicates an Infant School.
   I/J indicates a combined Infant and Junior School.
   F/M indicates a combined First and Middle School.
   Nursery School indicates a nursery with its own Headteacher.
   N indicates a nursery class attached to an Infant, First or Infant-Junior School.

5. Rural F (0): this is a rural First School which is not in a locality marked by social disadvantage.
Urban Inf. (disadvantaged): this school is in a locality marked by social disadvantage. (See Chapter 9 section 11 and Chapter 10.)

6. The school identification number is in parenthesis.

7. If assessments of achievement were requested from the school, the information is entered as +, Av or - in the following order: spoken language, mathematics (counting and computation), mathematical concepts, reading. 0 indicates a blank on the returned form. (See Chapter 8, iii, section 3 and Chapter 9, section 9) Teachers' written comments are in parenthesis.

3:6 Boy (863); morning N; 1st half-year; BTBC raw score 17; T score 52; passed items 3-6, 11, 13-16, 24, 28, 38, 40, 42, 44, 45, 47.
SES ?, 6. School (38): rural F (0).

3:6 Boy (720); morning N; 1st half-year; BTBC raw score 9; T score 40; passed items 2, 4, 5, 10, 11, 16, 20, 22, 40.
SES ?, 6. School (33): rural F (disadvantaged).

3:6 Girl (1255); full-time N; 1st half-year; BTBC raw score 14; T score 48; passed items 2, 4, 5, 8, 11-13, 16, 19, 21, 24, 28, 29, 34.
SES V, 3: M unemployed (SPF). School (49): inner city F.

3:6 Boy (1773); afternoon N; 1st half-year; BTBC raw score 12; T score 45; passed items 1, 2, 5, 10, 11, 12, 16, 21, 24, 30, 31, 40.
SES ?, 6. School (45): rural I/J (0).

3:6 Boy (1768); morning N; 1st half-year; BTBC raw score 14; T score 48; passed items 1, 2, 5, 10, 13, 14, 16, 19, 20, 21, 30, 31, 32, 34.
SES III SM, 2: F miner, M housewife. School (45): rural I/J (0).

4:0 Boy (395); morning N; 2nd half-year; BTBC raw score 20; T score 46; passed items 2, 4, 5, 7, 10, 13-16, 24, 26, 28, 30, 32, 34, 39-42, 44.
SES III SM, 2: F mechanic, M housewife. School (20): rural F (0).
Av (rather insecure - tearful); +; +; +.

4:0 Boy (715); morning N; 1st half-year; BTBC raw score 24; T score 51; passed items 1, 2, 4-7, 9, 11, 16, 20, 21, 23-26, 28-33, 36, 39, 43.
SES III WC, 2: F sales rep., M housewife. School (33): rural F (disadvantaged).

4:0 Boy (727); afternoon N; 1st-half; BTBC raw score 4; T score 26; passed items 1, 2, 5, 16, 20. Except for items 1 and 16, he always pointed to the first drawing. So the test was stopped after item 22. It was judged that item 20 was definitely passed by chance and items 2 and 5 may have been also.
SES V, 3: F unemployed, M housewife. School (33): rural F disadvantaged).

4:0 Girl (862); morning N; 1st half-year; BTBC raw score 26; T score 54; passed items 1-6, 8, 10, 11, 13, 15, 16, 19, 21, 22, 24, 26, 27, 30-34, 40, 42, 45.
SES III WC, 2: F policeman, M housewife. School (38): rural F (0).

4:0 Boy (1233); afternoon N; 1st half-year; BTBC raw score 20; T score 46; passed items 1-5, 7-11, 13, 16, 19-21, 24, 25, 28, 40, 41.
SES III, 2: F driver, M housewife. School (48): inner city Nursery School.

4:6 Boy (449); morning N; 2nd half-year; BTBC raw score 18, T score 39; passed items 1, 4, 6, 7, 10, 11, 13, 14, 16, 17, 19, 21, 24, 25, 34, 35, 40, 50.
SES V, 3: F unemployed, M housewife. School (21): urban F (0) Av; 0; +; Av.

4:6 Boy (842); reception class; 1st half-year; BTBC raw score 38; T score 64; passed items 1-12, 14-16, 18-22, 24-27, 30, 31, 33, 34, 36, 39-46, 48.
SES I or II, 1: F systemic analyst, M housewife. School (37): urban F (0).

4:6 Boy (1863); morning N; 2nd half-year; BTBC raw score 39; T score 66; passed items 1-22, 24, 27, 28, 30, 32-37, 39-41, 43, 48-50.
SES III SM, 2: F miner, M housewife. School (40): rural F (0).

4:6 Girl (1166); morning N; 1st half-year; BTBC raw score 37; T score 62; passed items 1-4, 6-9, 13-16, 18-24, 30, 31, 35, 36, 38, 40, 42, 44.
SES V, 3: M unemployed (SPF) School (46): rural Inf., CE (0).

4:6 Girl (1176); afternoon N; 1st half-year; BTBC raw score 23; T score 45; passed items 1-8, 10-13, 16, 18, 20-22, 24-26, 39, 40, 41.
SES ?, 6. School (46) rural Inf., CE (0).

5:0 Boy (1180); reception class; 1st half-year; BTBC raw score 24; T score 41; passed items 1, 2, 4, 5, 7-11, 13-16, 19, 22, 24, 26, 34-36, 40, 42, 43, 45.
SES ?, 6: Father works at Iron Co., M housewife. School (46): rural Inf., CE (0).

5:0 Girl (1286); reception class; 1st half-year; BTBC raw score 34; T score 52; passed items 1-4, 7, 9-25, 27-31, 34-36, 38, 40-42.
SES I, 1: F Dr., M housewife. School (51): urban F (0).

5:0 Boy (918); reception class; 1st half-year; BTBC raw score 40; T score 63; passed items 1-19, 21-30, 32-34, 36-40, 42, 48, 49.
SES IV, 3: F process worker at Coalite plant, M housewife.
School (40): rural F (0).

5:0 Girl (798); reception class; 1st half-year; BTBC raw score 40; T score 63; passed items 1-60, 18, 20-25, 27-30, 33-38, 40-43, 46-48.
SES III SM, 2: F bricklayer, M housewife. School (36): urban F (0).

5:0 Girl (654); identical twin; reception class; 1st half-year; BTBC raw score 36; T score 55; passed items 1-14, 16-25, 27, 30, 32, 34-36, 40, 41, 43, 47, 48, 50.
SES III SM, 2: F fitter, M housewife. School (30): urban F (0).

5:6 Boy (1260); class 2; 1st half-year; BTBC raw score 41; T score 59; passed items 1-7, 9-18, 20-28, 30-32, 35, 37-44, 47, 49, 50.
SES I, 1: F Dr., M Dr. Live in hospital flat. Asian. School (49): inner city F.

5:6 Girl (1374); class 6; 2nd half-year; BTBC raw score 38; T score 54; passed items 1-4, 6-11, 13-21, 23-25, 27-30, 33-36, 39-43, 46, 47, 50.
SES III or IV/III WC, 2: F baker, M Post Office clerk. School (54): urban F (disadvantaged).

5:6 Boy (726); class 7; 1st half-year; BTBC raw score 30; T score 44; passed items 1, 2, 4-21, 24, 26, 27, 32, 35, 38, 40, 42, 47, 48.
SES V, 3: M unemployed (SPF). School (33): rural F (disadvantaged).

5:6 Girl (54); reception class; 2nd half-year; BTBC raw score 36; T score 52; passed items 1-11, 13, 14, 16-18, 20-22, 24, 25, 27-30, 32, 34, 35, 38-40, 42, 43, 46, 47, 49.
SES V, 3: F unemployed, M unemployed. School (4): rural I/J (0). School returned assessment forms but no assessments given for this child or for some other children in her class.

5:6 Girl (835); class 2; 1st half-year; BTBC raw score 25; T score 37; passed items 1, 3-5, 7, 8, 10-12, 16, 18, 20, 23, 24, 25, 27, 30-34, 38, 40, 41, 43.
SES V, 3: no father; mother and child live with grandmother who is child's official guardian. Mother is ESN (S). School (37): urban F (0).

6:0 Girl (1162); middle-infant class, 1st half-year; BTBC raw score 37; T score 46; passed items 1-14, 16, 17, 18, 20, 21, 24, 27, 29-35, 37, 39-44, 47, 48.
SES II, 1: F steel engineer (highly qualified), M housewife. School (46): rural Inf., CE (0).

6:0 Girl (177); middle-infant class; 2nd half-year; BTBC raw score 36; T score 44; passed items 1-7, 10, 11, 13-24, 27, 29, 30, 32-35, 37, 38, 40-43, 46, 49.
SES III, 2: F Senior foreman, BSC, M housewife. School (11): urban Inf. (0). Av; Av; Av; Av.

6:0 Girl (1280); middle-infant class; 1st half-year; BTBC raw score 31; T score 37; passed items 1, 2, 4, 5, 7, 10-19, 22-28, 30, 31, 33-36, 40, 42, 47.
SES V, 3: M at home, pregnant (SPF, divorced). School (50): urban F (disadvantaged).

6:0 Girl (1547); class 4; 2nd half-year; BTBC raw score 45; T score 63; passed items 1-13, 15-25, 27, 29-36, 38-42, 44-50.
SES V, 3: F unemployed, M unemployed. School (64): inner city F

6:0 Boy (989); middle-infant class; 1st half-year; BTBC raw score 38; T score 48; passed items 1-24, 27, 28, 30, 32-24, 36, 38, 40-43, 47, 48.
SES III, IV or V, 4: F occupation not known but not professional or managerial, M part-time cleaner. Jehovah Witness with consequent restrictions on child. School (28): rural F, CE (0) though some children travel from a disadvantaged estate.

6:6 Girl (1681); middle-infant class; 2nd half-year; BTBC raw score 44; T score 54; passed items 1-7, 9-18, 20-25, 27, 29-44, 46-49.
SES III, 2: F driver, M housewife. School (71): urban Inf. (0) in seaside town.

6:6 Boy (518); middle-infant class; 1st half-year; BTBC raw score 41; T score 48, passed items 1-27, 29, 30, 33, 34, 37, 39-42, 46-50.
SES II, 1: M social admin. officer, DHSS (SPF). School (24): urban F, CE (0). Av; Av; 0; Av.

6:6 Girl (747); middle-infant class; 1st half-year; BTBC raw score 38; T score 43; passed items 1-22, 24, 25, 27-32, 40, 42, 43, 45, 46, 48, 49.
SES V, 3: F third surrogate father, occupation unknown but manual, M unemployed. School (33): rural F (disadvantaged).

6:6 Girl (545); class 3; 1st half-year; BTBC raw score 46; T score 59; passed items 1-25, 27, 29-32, 34-43, 45-50.
SES II, 1: F oil rig supervisor, M housewife. School (25): rural F/M. +; +; Av; +

6:6 Boy (1777); class 3; 2nd half-year; BTBC raw score 44; T score 54; passed items 1-24, 27, 29-34, 36-44, 46-49.
SES V, 3: M unemployed (SPF). School (75): urban F (0) coastal town.

7:0 Girl (202); class 3; 2nd half-year; BTBC raw score 41; T score 44; passed items 1-11, 13-25, 27-34, 37-41, 43, 44, 46, 47.
SES III, 2: F fitter, M housewife. School (11): urban Inf. (0).

7:0 Girl (564); class 3; 1st half-year; BTBC raw score 43; T score 48; passed items 1-4, 6-18, 20-30, 32-36, 38, 40-43, 45, 46, 48-50.
SES IV or V, 3: F labourer at factory, M packer. School (29): urban F (disadvantaged). Av; Av; Av; Av.

7:0 Girl (1004); class 3; 1st half-year; BTBC raw score 46; T score 56; passed items 1-7, 9-35, 37-48.
SES III WC, 2: F clerk, M housewife. School (41W): urban F (0).

7:0 Boy (108); class 2; 2nd half-year; BTBC raw score 47; T score 59; passed all items except 23, 33 and 44.
SES II, 1: F farmer, M nurse. School (7): rural I/J, CE (0).
N on school roll: 119; 4 classes; all 50 infants in one room with 2 teachers (old building). No assessment forms were returned by this school.

7:0 Boy (1009); class 3; 1st half-year; BTBC raw score 40; T score 42; passed items 1-7, 9-22, 24-30, 33-36, 39-43, 47, 48, 50.
SES II WC, 2: F salesman, M housewife. School (41W): urban F (0).

7:6 Girl (277); class 3; 2nd half-year; BTBC raw score 46; T score 53; passed all items except 26, 29, 38 and 44.
SES ?, 6: parents' occupation(s) not known. School (6): rural I/J (O). No assessment forms were returned by this school.

7:6 Boy (620); 7/8 class; 1st half-year; BTBC raw score 41; T score 41; passed items 1-27, 30, 32, 34, 37-42, 44-46, 49, 50.
SES ?, 6: parents' occupation(s) not known. School (31): urban F (disadvantaged).

7:6 Girl (1397); class 5; 2nd half-year; BTBC raw score 38; T score 36; passed items 1-9, 11-25, 27, 30, 32, 38-42, 44, 46-49.
SES III WC, 2: F Country and Western musician, M housewife. School (55): urban F (disadvantaged).

7:6 Girl (514); 1st year Junior class; 1st half-year; BTBC raw score 47; T score 56; passed all items except 36, 38 and 44.
SES II, 1: F C.C. admin. officer, M housewife. School (24): urban I/J CE. + (Spoken language very good indeed); +; + (good understanding of all mathematical concepts); + (fluent reader - reads with understanding).

7:6 Boy (998); class 7; 1st half-year; BTBC raw score 41; T score 41; passed items 1-13, 15-18, 20-29, 32, 24, 25, 38-46, 48, 49.
SES III, 2: F crane driver, M housewife. Child has speech therapy. School (35): urban F (O).

Appendix 16

Errors on each item are located in the tables as follows:

Item	Tables	Pages
1	57, 58	284, 285
2	59, 60	287, 288
3	59, 60	287, 288
4	59, 60	287, 288
5	58, 59	285, 287
6	51, 52, 53, 57	278, 279, 284
7	57, 58	284, 285
8	51, 52, 53	278, 279
9	51, 56	278, 283
10	58, 59	285, 287
11	59, 60	287, 288
12	58, 59	285, 287
13	57, 58	284, 285
14	57, 58	284, 285
15	57, 60	284, 288
16	59, 60	287, 288
17	51, 53, 57, 58	278, 279, 284, 285
18	59, 60	287, 288
19	57, 60	284, 288
20	29, 60	287, 288
21	52, 57	278, 284
22	57, 58	284, 285
23	58, 59	285, 287
24	51, 52, 57	278, 284
25	51, 60	278, 288
26	55	282
27	59, 60	287, 288
28	59, 60	287, 288
29	55	282
30	57, 58	284, 285

Item	Tables	Pages
31	57, 58	284, 285
32	57, 58	284, 285
33	59, 60	287, 288
34	54, 60	281, 288
35	57, 58	284, 285
36	57, 58	284, 285
37	51, 52	278
38	57, 58	284, 285
39	57, 58	284, 285
40	51, 52	278
41	54, 60	281, 288
42	57, 58	284, 285
43	57, 58	284, 285
44	59, 60	287, 288
45	51, 52, 57, 58	278, 284, 285
46	59, 60	287, 288
47	59, 60	287, 288
48	58, 59	285, 287
49	57, 58	284, 285
50	51, 52	278

Appendix 17

The raw data is too bulky to include in a bound copy. It may be obtained from the author c/o The Principal, Doncaster Metropolitan Institute of Higher Education, Waterdale, Doncaster DN1 3EX. (The Principal will forward correspondence.)