FROM CLINICAL STUDY TO STATISTICAL VERIFICATION:

A FURTHER STUDY OF CHILDREN'S ERRORS IN
THREE-TERM SERIES PROBLEMS

by

John H. Duthie, M.A.

Thesis presented for the Degree of Doctor of Philosophy
of the University of Edinburgh in the
Faculty of Social Sciences, January, 1965
I should like to express my gratitude to my supervisors, Professor James Drever and Professor John G. Pilley; and to Dr Margaret Donaldson, all of whom have read and criticised the manuscript at its various stages and many of whose suggestions have been incorporated into the text. Dr Donaldson has in addition been closely associated with much of the clinical work described in the pages which follow.

I should also like to thank Mr A.E.G. Pilliner for invaluable advice on the design of the experiments and on the statistical interpretation of the results. Mr R.B. Bett and Mr H.S. McNair have read and offered helpful criticism of several parts of the manuscript; and Miss Norma Bendell assisted with the clinical interviews of children in the study of partially quantified non-overlap problems. I am very grateful for their assistance.

For their co-operation my thanks are due especially to Miss Brown and the teachers of Moray House Demonstration School, and to the Directors of Education, the Education Committees and the teachers of the City of Edinburgh, East Lothian, Roxburghshire, and the City of Preston; and to the pupils of these schools who so willingly co-operated in the enquiry.
CONTENTS

Chapter 1: Introduction ........................................... 1

Chapter 2: Three-term Series Problems; Overlap Error: Clinical Findings .......... 20

Chapter 3: Overlap Error: Experimental Design and Group Test Results ............ 61

Chapter 4: Asymmetry Error: Clinical Findings, Experimental Design and Results .... 81

Chapter 5: Further Research on Overlap Error and Asymmetry Error: Validational and Longitudinal Designs and Results ................. 102

Chapter 6: Other Analytical Categories ............... 109

Chapter 7: Conclusions and Discussion ................. 111

Appendix I Supplementary Experimental Findings and Discussion ..................... 146

Appendix II Details of Data and of Statistical Calculations ............................. 1

Appendix III Problem-forms used in Group Tests ..........................xiii

References
CHAPTER 1

Introduction

I want to begin by examining the role played by method in the evolution of psychology, particularly in the psychological investigation of cognition, since it is with cognition, and especially with problem-solving, that the present study is concerned.

Questions of method are important in any science and psychology is no exception. Because of the relative immaturity of his study and the complexity of his subject-matter, the question of method is perhaps more explicitly considered by the psychologist than by any other kind of scientist. It was through the shortcomings of their methods, for example, that the Introspectionists lost favour to the Behaviourists and others: when the study of conscious content was abandoned in the second decade of the twentieth century, it was mainly because of a belief that no adequate methods were available for its controlled study. Of course the dogmatic Behaviourists went further, and, because they had no tools available for the study of consciousness, rejected consciousness as well as its methods of study. In 'Psychology from the Standpoint of a Behaviorist' (1924) for example, Watson wrote: "Even if they \[\text{people's consciousnesses}\] existed, they would exist as isolated, unusable 'mental' curiosities. The behaviorist finds no\footnote{my italics} evidence for /
for 'mental existences' or 'mental processes' of any kind.

In spite of the Classical Behaviourist criticism it is of course true to say that one of the most characteristic attributes of man is that he is conscious; that he is capable of experiencing the world around him. We know this both subjectively and also as a result of entering imaginatively into the experience of others and through communication with them. (See for example the writings of Buber (1942) and George Herbert Meade (1934).)

It is little wonder therefore that in its first flush of enthusiasm, experimental psychology attempted to take the bull by the horns and to tackle directly the problem of consciousness. There was no way of knowing then that the undertaking was doomed to failure (psychologically speaking) because of difficulties in method, and there could have been no way of discovering these difficulties in method other than by trying them out. These attempts thus appear, in retrospect, as an important first step in the psychological study of cognition.

In this study, which reached its peak under Titchener, a pupil of Wundt, content (consciousness) was primary and dictated method. As Boring (1950, P.385) put it: "It was introspective because consciousness was its subject matter." But, as I have suggested, it was because of its method that structural psychology went into a sharp decline. What was wrong with its method? Osgood (1953) lists four faults:

"(1)/

---

1 Of course I do not intend to tar all behaviourists with this brush. As I point out later, contemporary behaviourism has moved away from the position indicated here. This is perhaps most evident in the writings of Brunswik, who in his Probabilistic Functionalism, deals fairly directly with cognitive as well as perceptual functions. (See, for example, the chapter in Koch (1959a), in which Postman and Tolman state - P.510 - "In both these developments - the study of perceptual constancies and molar behaviorism - the emphasis shifts away from peripheral events to organismic achievement."
"(1) The data it yields are patently unverifiable.  
(2) Many relevant data are unavailable to the method.  
(3) Language is not a mirror of thought.  
(4) Only the effects of thought, not the process itself, can be observed."

It is interesting and important to note that the last three of these criticisms are also potential criticisms of the 'thinking aloud' technique discussed below (see page 16) and these criticisms will be discussed there. Meanwhile it should be noted that the first criticism seems unanswerable and to be the only final reason for the demise of Introspectionism. For this reason the questions asked by cognitive studies today are different in kind from those asked by Titchener: Titchener and his associates asked questions of a phenomenological or experimental kind; e.g. What is the quality of blackness? Today's questions, on the other hand, although often structural, deal not so much with the structure of subjective experience per se as with cognitive structures as postulated by others, and eventually (if Hebb, 1949, is correct in his predictions) as observed by others (see below).

The methodological restrictions imposed by the rise of Behaviourism (and also those imposed by the parallel Gestalt movement, although these were less severe - see below) led to the eclipse of anything which might be called the study of the cognitive processes for the next two decades. I am speaking here of psychology in the United States and in Gt. Britain. In Europe, which the positivist movement has virtually passed by, there was of course a different history. And there too, Jean Piaget had already started his study of cognitive processes, much to the chagrin of those scientific psychologists in the U.S.A. and in Gt. Britain who recorded the fact /
fact that they noticed (see e.g. Isaacs, 1929; Curti, 1938; Hazlitt, 1930).

The Behaviourists hoped that by a study of stimulus and response, and the use of statistical techniques, they might be able adequately to describe, explain and predict some, or all, of the behaviour of man. Watson, himself, however, was not absolutely clear about the potential relative importance of method and of 'system' for Behaviourism. In 1924 he wrote, "Whether it is to become a dominant system of psychology or to remain merely a methodological approach is still not decided" (P.vii). On the question of system, he wrote, "Such a clean break with the whole concept of consciousness is possible because the metaphysical premisses of Behaviourism are different from those of structural psychology" (P.viii), showing that although he was shrewd enough to see the value of Behaviourism purely as a method, he himself saw its chief merit in its role of philosophical system.

Whether we take Watson's contribution as system or method however his model for the study of behaviour can be described as S - R and hence reactive (see pages 6 ff.). And again it can be claimed that such a model was a necessary step in our attempt to find a method appropriate to the study of behaviour and cognition. But it gradually became apparent that the

Of course Watson did not create this revolution in thinking alone any more than did Titchener create Introspectionism. But all of Watson's predecessors had been unwilling or unable to take the final logical step along the road to Behaviourism - the complete rejection of consciousness as a subject for psychological study. Watson accused Thorndike, who introduced S - R methodology, and Bechterev and Pavlov from whom he borrowed the techniques and principles of conditioning, quite explicitly of this failure; e.g. "Those so-called objectivists so far as concerns their human psychology - and this is true of Bechterev as well - are perfectly orthodox parallelists. While behaviour psychology borrows the conditioned reflex methods from Pavlov and Bechterev, and its courage to face all aspects of human behaviour from the psychopathologists it is neither an objective psychology in Bechterev's sense nor a modified system of psychoanalysis" (Pp.xi-xii). The full and final rejection of everything Introspective-structuralism stood for may thus be seen as coming only with Watson.
S - R model was inadequate. Something was going on between stimulus and response which made it impossible to predict response from a knowledge of stimulus alone. This growing realisation about the need for central processes resulted from the application of S - R methods. As Hunt (1961) paradoxically puts it, "... stimulus-response methodology... has been rediscovering 'mind' ... stimulus-response methodology has been undoing stimulus-response theory."

The first evidence against S - R theory came, strangely enough, before the publication of Watson's theory. In a paper published in 1912, W.S. Hunter, as a result of experiments on delayed reaction, argued that animal behaviour could only be explained on the basis of some form of symbolic process. He strengthened his case in 1918 by the use of the double-alternation maze, a form of temporal maze in which a right turn can signal either a right or left turn, and which is therefore difficult to account for in S - R terms.

Attempts to overcome difficulties such as these resulted in the introduction of intervening variables and hypothetical constructs, tied at both ends to S and R. The organism was still conceived of as reactive rather than active or self-organising (see pages 6 ff.). The mediational process can be conceived of rather as a resistor or condenser in an electrical circuit whose value is given by the past experience of the organism: $S \rightarrow \rightarrow R$.

This diagrammatic representation of the postulated process is of course in no way conceived as directly descriptive (see below). Rather is it intended to emphasise the relation between the new theorising and its roots in classic Behaviourism.

In/
In a sense, intervening processes of this kind were introduced by Watson himself at an early stage when he suggested that thinking was sub-vocal speech (e.g. Watson, 1924, P.14, although as we have seen, the first introduction came before the publication of his first paper.) The most notable of the attempts to formulate intervening variables precisely was of course that of Hull (e.g. 1943), whose conception of 'pure stimulus acts' was seminal (e.g. in the Mediation Hypothesis of Osgood - see below).

It is rather difficult to see the precise relation between attempts such as those of Hull to construct intervening variables, and those of for example Tolman (who, incidentally, introduced the term in 1932). Hunt (1961, P.68) and Osgood (1953, Pp.653-654) see the distinction in terms of a peripheral-central dichotomy. This dichotomy appears to refer to the postulated locus of the hypothesised activity, i.e. whether it is within the central nervous system or in the afferent or efferent systems. Here is how Osgood puts it: "The real issue is the locus of the critical nervous activity..." (P.653). Now this distinction, although relevant, does not seem to me to be critical. Another, and this time critical, distinction seems to be among three possible models: the reactive model in which the organism is seen as responding to a stimulus and otherwise (in the absence of stimulation) failing to respond at all (e.g. Hullian theory); the active model in which the organism emits behaviour 'spontaneously' but in which the learned activities of the organism are just as much the result of a simple relationship with the environment as in the case of the reactivists (e.g. Skinner's system); and the self-organising model in which the organism is seen not only as active in the environment, but also as processing the input and reacting or not as a result of this processing (e.g. Hebb's autonomous central processes).
Returning to the peripheral-central dichotomy, it seems that if a locus is postulated at all then according to MacCorquodale and Meehl's distinction (1948) the theorist is employing a hypothetical construct. The alternative, in their terminology, is of course the 'intervening variable', a postulate which says nothing concerning locus and which describes a relationship between two variables, often in mathematical terms, involving no surplus meaning (see also page 13, footnote 1).

It seems that the differences among the various 'schools' can in fact be described by creating a 3x3 matrix from these sets of distinctions:

<table>
<thead>
<tr>
<th>Hypothetical Construct</th>
<th>Intervening Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>locus</td>
</tr>
<tr>
<td></td>
<td>peripheral</td>
</tr>
<tr>
<td>active and self-organising</td>
<td>-</td>
</tr>
<tr>
<td>active but not self-organising</td>
<td>-</td>
</tr>
<tr>
<td>reactive</td>
<td>Hull e.g.</td>
</tr>
<tr>
<td></td>
<td>rg</td>
</tr>
</tbody>
</table>

Examples have been suggested to show how the matrix can be given concrete reference to the history of psychology. The reason for some of these entries (e.g. Tolman who as I have said introduced the term 'intervening variable') is self-evident, although it should be noted that these entries are 'best fits' rather than exclusive classifications. MacCorquodale and Meehl (1948), for example, say that Tolman is not always consistent in his use of intervening variables and that even Skinner appears to slip /
slip over into hypothetical construct on one occasion, when he talks of 'the strain on the reserve'. It might be as well, however, to say a little about the reasons for other entries. Hull is placed in two categories according to a distinction made by MacCorquodale and Mehl who classified Hull's theory as utilising both hypothetical constructs and intervening variables. Both Piaget and Hebb - to anticipate a later discussion - occupy the same cell because of similarities in their theories of adult cognitive functioning. There is an interesting distinction, however, in their early learning, Hebb's being much more perceptually-oriented and Piaget's behaviourally oriented. Even the use of the reflex, according to Piaget (1953b) is active in the first stage of development (e.g. p.25: "Now this sucking schema... is not limited to functioning under compulsion by a fixed excitant, external or internal, but functions in a way for itself.") Piaget appears in this way to provide a corrective to Hebb's reactive model for early behaviour. As Drever (1964) says, "The principal advantage of starting with motor activity... is that we may find it possible to substitute known reflexes for hypothetical 'unities and identities'."

What I am saying then is that the importance of the systems of intervening variables suggested by such men as Tolman (e.g. 1932, 1937, 1948) and Krechevesky (1932) lies in the fact that these represent the first faltering attempts to free the organism from its slavery to stimulus and response. These attempts were rightly criticised at the time (e.g. Spence, 1942) as failing to explain "... the process of inventive ideation..." (ibid. p.297), but in the light of subsequent happenings (principally Hebb's and Harlow's 1949 publications) they appear as perceptive attempts to signpost the direction for future research.
One other (and probably the last) major attempt made to construct intervening variables of the reactive sort was that of Osgood (1953, p.392ff), who in his Mediation Hypothesis attempted to tie central processes in the form of meditational responses firmly to externally manipulable stimuli. Although he manages in this way to introduce such organism-active processes as attention such processes are still seen as resulting from stimulation by a 'sign'. Taken as a molar unit, it seems that the organism was still conceived as reactive. The evidence against such a position will be reviewed in the description of Hebb's work below.

Meanwhile it is necessary to note the contribution of another school of psychology, contemporaneous with the school of neo-Behaviourism which has just been described; I refer of course to Gestalt psychology. Gestalt psychology, which at the time was seen to be in direct opposition to the Behaviourist approach, was almost as stimulus-bound as Behaviourism (although they would never have agreed to the term 'stimulus'). Their modification at the stimulus end (a relatively slight one in view of the revolution to come) was in terms of patterns of input. (Their chief objection to Structuralism had been in terms of its atomism.) To the extent that they introduced the concept of insight they may be said to have paved the way for a return to the study of cognition; but their contribution was, in retrospect, a slight one, overemphasising as it did innate cortical patterning at the expense of learning (I am speaking here of the typical Gestalt study such as Kohler's, 1925; there were exceptions, witness the work of Tolman already referred to); and disproved as the nativist view was by subsequent experiments such as Harlow's work on learning sets, 1949. Methodologically, they toed the Behaviourist party line, reacting in much the same way to Introspectionism.
The revolution, whose first rumblings we noticed above, came in 1949 with the publication of Hebb's historic text, "The Organization of Behavior". And again the revolution can be viewed, at least in part, as a methodological one. Almost by definition, one of the tenets of S-R methodology was that one should remain outside the organism. It is true, as we have seen, that processes intermediate between S and R were soon considered necessary, but these, more often than not, were considered to have a peripheral rather than a central locus, and thus were safely outside the organism too. In the Hullian reactive organism, the stimulus still was king (and the response his queen).

Hebb's revolution was two-fold. His first move was to show how set, attention and autonomous central processes generally (and Gibson's 1941 paper had shown just how general such constructs were) could be adequately accounted for without reintroducing a ghost into the machine. And he did this by postulating feasible physiological mechanisms, well-rooted in contemporary neurophysiological knowledge. (As Hebb himself admits he was not the first to link possible central processes with physiology. Papers such as Denny-Brown's (1932) had prompted Meehl and MacCorquodale to make their theoretical distinction between intervening variable and hypothetical construct a year before the publication of Hebb's text.)

What I am saying is that Hebb suggests as a method that we should subject our psychological theorizing to the requirement that, in principle at least, it should be realisable in terms of neurophysiology (although a detailed specification in terms of locus etc. is not necessary, at least initially - see 1958, Pp. 460 and 462); or in other words, that we should attempt /
attempt to build neurophysiological models for our psychological theories. And this leads us to his second contribution, which was to show how the Hullian approach was at variance with the neurophysiological evidence.

If we take account of the neurophysiological evidence, he says, then "the conclusion becomes inevitable that the non-sensory factor in cerebral action must be more consistently present and of more dominating importance than reluctant psychological theory has ever recognized" (1949, P7). The neurophysiological evidence which, according to Hebb, made S - R theory finally untenable falls into two kinds. Firstly, "...the central nervous system is continuously active, in all its parts, whether exposed to afferent stimulation or not" (P8). Advances in neurophysiological knowledge since the publication of Hebb's theory have according to Pribram (1960) shown the brain to be even more active than Hebb conceived it to be. And secondly, Hebb claimed that the evidence was now firmly against the original physiological hypothesis that neural transmission is simply linear; and again, subsequent evidence both from neurological and psychological sources has amply confirmed his original statement (see, for example, the review by Melzack (1961) in which he states that "there is no longer any doubt that these message-modifying fibres exist; it has been found that electrical stimulation of widespread regions of the brain is able to modify the messages transmitted through every major sensory system" (P6). Perhaps the most striking psychological evidence is provided by the experiments of Thompson and Melzack (1956) in which Scottish terriers, raised in isolation, failed to react (other than by sniffing) when a burning match was held to their noses.)
It is interesting to note that Hebb does not merely say that his approach is one possibility; he sees it as ousting "mathematical models and factor analysis... The long and short of it is that you cannot get out of an equation, or any mathematical gymnastics (would he include computer simulation?) any better ideas than you put in..." (ibid, P455).

Fifteen years later we are still building upon and integrating into the general framework of psychology Hebb's ideas as contained in his 1949 text. Two examples of the current worth of Hebb's 1949 theorising should suffice.

An important part of Hebb's theory was the role of sensory input for adaptive behaviour. (Although he did not go along with the stimulus-determined behaviour of the Behaviourists, neither did he see the stimulus as unimportant, as did some cognitive theorists, notably Lewin (e.g. 1935).) "It does seem clear from the facts discussed, from the large potentials observed in sleep, and from the hypnotic effect of minimising the normal variation of sensory activity, that the sensory input to the brain has a constantly necessary function, for adaptive behaviour". (1949, P10). Hebb was not the first to recognise such effects, but he was the first to grasp their significance, and his treatment of them gives some theoretical orientation to subsequent studies on sensory deprivation (see e.g. Solomon (1961)), although he himself has since claimed (1961) that an adequate theory of stimulus deprivation is still lacking. (Subsequent attempts at such a theory, together with a review of further studies, are contained in Zubek (1964).)

Hebb himself in Koch (1959a, P637) gives a brief review of recent work which he considers has been stimulated by his theorising. He also sees his theory as having raised "...rather general questions and promoting argument in the laboratory" (ibid).
Computer simulation of human behaviour (e.g. Newell, Shaw and Simon, 1957, and the various papers contained in Tomkins and Messick, 1963) has shown if unwittingly how well Hebb's autonomous central processes can be realised in 'hardware' as well as in 'wetware' and since these constructs have considerable surplus meaning\(^1\) one wonders whether Hebb would accept the value of such models in place of, or in addition to, physiological ones. Certainly they do not appear to be open to the same objections that he makes to mathematical models (see above). (Messick, 1963, p.309, makes the point that the use of computer models also results in an increase in rapidity and also in the complexity of the relationships which can be handled.) Computer simulation studies appear to accord well with Hebb's general position and can be considered a vindication of it (see below).

Interestingly these computer models have provided students of cognition with a new set of terms and concepts - Miller, Galanter and Pribram (1959) for example have suggested replacing the reflex-arc with the TOTE unit as the basic unit for behavioural study, a concept derived directly from computer simulation studies - thus forming a nice feedback loop for the modification and improvement of the study of cognition\(^2\).

---

1 Colby, 1963, for example, states that "...it is incorrect to say that nothing new can come from a programme because we know everything that was put into it initially. One may know all the initial data, but in a developing system evolving over time which can create an unlimited number of derivatives, we cannot know all the implications and consequences of what we have put in. A programme generates many surprises, even for the programmer" (p.177). 'Surplus meaning' was an expression used by Maccormudale and Meehl as one of the ways in which 'hypothetical constructs' can be distinguished from 'intervening variables'.

2 I do not mean to suggest, of course, that all investigators have gone over to the active self-organising organism approach. Hebb has pointed out that although "...the day of cut-and-cut peripheralism is over, there is still a certain cramping effect of older ideas" (1958, pp.457-458). A recent issue of Monographs of the Society for Research in Child Development (Kessen and Kuhlman, 1962) for example, devoted to a study of the work of Piaget, took a predominantly SR approach and attempted to account for Piaget's findings in these terms. Papers in the report of the second conference (Wright and Kagan, 1963) took a similar organism-reactive line.
Most of the studies which have so far been described can, as I have suggested earlier, be contrasted with the studies which were carried out in Europe during the same period. The distinction can again be seen as a methodological one, the European investigators employing a clinical method (see below) in contrast to the more statistically oriented studies of the American and British investigators.

To date, each group has tended to distrust the other. The clinician has believed (on the whole, correctly, I think) that human behaviour is too complex for the *a priori* hypotheses of the statistician to be fruitful; and the statistician, although he has not denied that the individual case may be lawful, has believed (again, it seems, rightly) that human behaviour is too complex for the clinician to be certain of his findings.

The last ten years or so have however seen the beginnings of a rapprochement between the two camps, each admitting the merits of the other, if not their own limitations. Thus a new interest is being shown in clinical work in the U.S.A.: witness the texts by J. Mc.V. Hunt (1961) and by J.H. Flavell (1963), describing in detail the work of Piaget and also the American conferences to discuss his work, the latest of which, at Cornell, in March 1964, was addressed by Piaget himself (transcript to be published in *J. Res. in Sci. Teaching*). And in Europe, Piaget is at last recognizing the need for statistical verification of his clinical studies: in his foreword to Flavell's (1963) new book, Piaget writes: "As to the criticisms usually directed toward me regarding the qualitative method which we use in our intellectual development studies... it must again be said /
said that our research is far from completed and that all sorts of controls, both statistical and nonverbal are currently in progress."

(For a more detailed critique of Piaget's present position see Pp.130 ff.)

Even greater progress than this toward an adequate rapprochement has been made, however. Holt (1950) for example, has made specific suggestions as to how the rapprochement might best be made: "Rather than continuing to apply obviously inadequate statistical methods, clinical researchers might do much better to concentrate on intensive studies of single cases: observing in as controlled a way as possible, trying to discern meaningful relationships and to set up hypotheses which may be tested when appropriate methods for establishing proof are at hand. It is all too often forgotten that statistical methods are primarily ways of proving (or more exactly disproving) hypotheses and only secondarily means of finding something out... So great is the scientific glamour of more experimental methods that there is a danger that too little research of this phenomenological kind will be done."

The suggestion then is to unite the two defective methods (clinical and statistical) each defective for the reasons given by its opponents, into a single effective one: the use of the clinical method to try to derive fruitful hypotheses, and the use of statistical methods to check the supposed relationship by isolating it statistically in a large number of instances.

Another area in which the need for a rapprochment between clinical and statistical method has been realised is, paradoxically, that of comparative psychology (see, e.g. Harlow, 1958, P.6; also Verplanck and Hayes,1' In/
In animal studies, automation had led to such a reliance on statistical results that there was a real danger of losing sight of the actual behaviour of the animal. Thus it was hypothesised for example that learning would be in direct ratio to the size of the reward until it was noticed that rats behaved differentially to different sizes of food pellet, consuming the smaller ones at once, and hoarding the larger ones for consumption later. According to Bett (personal communication) graduate students in Harlow's laboratories are initially advised to forget their preconceptions and to spend the first part of their studies simply observing animals until a hypothesis occurs to them as a result of such observation. (Hebb makes a strikingly similar point in another context, 1958, Pp.463-464.)

It was not until 1963 that a similar statement was made in the field of cognition, although when it did come, the statement was more specific and more detailed. Donaldson, whose statement it was, suggested that the most appropriate clinical method in this case was the 'thinking aloud' method. Four possible objections to this method have already been listed (P.3). The first objection does not apply to thinking aloud (although as I suggested above it is unanswerable in the case of introspection), when, as Donaldson suggests (ibid), the process is used only to generate hypotheses which must then be tested statistically. Osgood's second objection, that many relevant data are unavailable to the method, is only critical if we are naive enough to believe that 'thinking aloud' (or for that matter introspection) tells us all there is to know about thought. It is surely better to gain some knowledge in this way than none at all. (An interesting reason for the possible lack of availability of cognitive processes /
processes to introspection - and hence to thinking aloud - comes from the U.S.S.R.: Galperin (1954). He suggests that this method will reveal only processes which have not yet become habitual: "It has long been known that as we master a psychic process, it becomes automatic; in current terms it becomes a dynamic stereotype and 'drops out' of consciousness." (P213) And "...it loses the remnants of sense content, so that introspection cannot now reveal its actual course." (P222).

The solution which he suggests is the study of the formation of cognitive processes.) Osgood's third objection is rather more difficult to deal with (Language is not a mirror of thought) since it is not clear (at least to me) how this objection differs from the last one. Does he mean that language does not reflect thought at all? Hardly. Does he mean that language does not reflect everything that goes on in the mind? Then this objection is equivalent to the last one. If he means that the mirror may be a distorting one (and this is a legitimate criticism, although it is not clear that this is what he means) then statistical verification will take care of this danger too. That only the effects of thought can be observed (Osgood's fourth objection) is not a fatal criticism of the method since scientific observations have often successfully been made at this level, especially in the natural sciences - witness for example, the use of the cloud-chamber to study the effects of sub-atomic particles which were otherwise unobservable.

In her text, Donaldson did not in fact go beyond the stage of clinical study (nor has she in her subsequent work). The present study represents the first attempt in this series to verify statistically hypotheses which have been clinically derived.
In reporting the findings of the present investigation, two approaches seem possible — an historical approach (in which the work is reported in historical sequence); or a formal approach (in which evidence is considered out of historical context and the sequence of the original is lost). Since the work reported deals only with one kind of problem — the three-term series — it seems that an historical approach is possible without the repetition that would have been necessary had more than one problem been used. And an historical approach more accurately reflects the theoretical position adopted here. Nevertheless, there are occasions on which it is necessary to anticipate findings and hence to adopt a more formal method of reporting: without this the different emphases placed upon clinical findings (since some sort of selection is necessary in a report of this kind) might at the time seem arbitrary; again there are times when it becomes necessary to delay the statement of findings so that the description of events does not become unnecessarily complex; one other reason for the more formal arrangement which is occasionally adopted is the occurrence of two unrelated formal requirements within the three-term series problem: overlap and asymmetry. Since these do not seem to be related (other than by their chance occurrence within the same kind of problem) it has been thought better to deal with them in separate sections of the report. The orientation adopted is thus historical with occasional concessions toward a more formal approach.

The /

---

1 A report of the findings at the clinical level has been published by the present author as an appendix to Donaldson’s book, 'A Study of Children's Thinking' (1963).

2 These terms will be explained in detail later.
The report of the research thus falls into five main sections:

(a) A description (Chapter 2) of the clinical studies during which the nature of the structural errors involved in three-term series problems was studied, and during which an attempt was made to strip the problem-forms of ambiguity and unnecessary complexity.

(b) A description of the statistical investigation into overlap error based on this clinical stage and of the results obtained from it (Chapter 3).

(c) Clinical and statistical evidence on the nature of asymmetry error (Chapter 4).

(d) The experimental design and results of the validation and longitudinal studies of both the overlap and asymmetry errors (Chapter 5).

(e) A summary of the conclusions which can be drawn from the results of the investigation as a whole, together with a discussion of the implications of these results for the theoretical position described here (Chapter 7).
CHAPTER 2

Three-term Series Problems: Overlap Error: Clinical Findings

I have already stated that the present study represents the first attempt (at least in the present series) to test statistically hypotheses which have been clinically derived. The hypotheses tested in the present study concern two errors made by children in three-term series problems, the so-called overlap and asymmetry errors. The name 'three-term series', first proposed by Hunter in 1957, refers to a problem type which has often been used in the study of cognition (see e.g. Burt, 1919 and Piaget, 1921). Donaldson (1963) also used the problem, and defined it as "...any one in which, from the information provided, the subject must assign absolute or relative values to one or more of three objects which stand in serial relationship to one another" (ibid, P83). Donaldson's isolation (by clinical means) of the overlap error and the asymmetry error provided the jumping-off point for the present study which was aimed at the further experimental definition, both by clinical and statistical means, of these two errors - definitions which as we shall see are rather far removed from those postulated as a result of the initial clinical findings and which show clearly the need for the statistical verification of findings made initially at the clinical level.
Two kinds of three-term series problem are used in the present investigation, one of which has been termed 'fully quantified' and the other 'partially quantified'. In order best to illustrate the difference between fully quantified and partially quantified forms, there follows a fully quantified problem which was developed only as the result of the extensive clinical investigation to be described later:

Peter is 9.

Peter is three years older than Ian. Peter is two years older than David.

How old are Ian and David?

The partially quantified form is identical with the fully quantified form except that the information concerning Peter's age is omitted. The subject is instead required to find the size of the difference between the dependent variables:

Peter is three years older than Ian. Peter is two years older than David.

What is the difference between the ages of Ian and David?

(The need for, and the significance of the partially quantified form became clear only as the investigation proceeded.)

Now in both fully quantified and partially quantified problems (as indeed in all three-term series problems), one of the terms must be mentioned twice. Following Hunter, this term has been called the 'link'. In the problems mentioned so far, the link has been at the end of the series. As a result, in the problems as stated, the two relationships overlap /
overlap one another, and because of this, this form of the three-term series has been called the 'overlap form' (hereafter called simply, 'overlap problems'). For example, in the problem quoted above, the relationship Peter-Ian overlaps the relationship Peter-David. Another possibility exists: that in which the link is in the middle of the series. Both fully quantified and partially quantified forms of this type are also possible. The partially quantified form is given here as an example:

Betty is four years older than Susan. Susan is three years older than Ann.

What is the difference between the ages of Betty and Ann?

This form of the three-term series problem, called the non-overlap form, was used only much later in the investigation. Formally, however, we can say that there are four possibilities: fully quantified overlap/fully quantified non-overlap/partially quantified overlap/partially quantified non-overlap. (Another possibility, the unquantified problem - which is soluble only in its non-overlap form - was not studied in the present investigation, but has already been studied fairly extensively by Burt (1919), Piaget (1921), Hunter (1957a) and most recently by Witherington - see Donaldson's text - ibid, Pp115-133.) Other variations of these four types are given later, as they are required.

The overlap form of the three-term series problem was first used by Donaldson (ibid, Pp86-87) in the two versions given below:

Problem Bl: /
Problem B1:

We want to find out the ages of two girls called Jean and May. We know that a third girl, Betty is 15, and that she is 3 years older than one of the two girls and 5 years older than the other. If we had one more piece of information we could calculate the ages of Jean and May. What is that piece of information?

Problem B2:

Tom, Dick and Harry are 3 boys. Dick, who is 5 feet 4 inches tall, is 6 inches taller than one of the other boys and 2 inches taller than the remaining one. Harry is taller than Tom.

Therefore:

1. Tom is 5' 2" tall.
2. Harry is 4' 10" tall.
3. Harry is 5' 0" tall.
4. Tom is 4' 10" tall.
5. Harry is 5' 2" tall.

Although it is the purpose of the present study to provide experimental definition, some preliminary (and at this stage necessarily superficial) definition of the first of the errors to be studied is in order. This can perhaps best be provided by giving concrete examples. In the first problem stated above (P21) for example, the overlap error would yield the answers 'Ian is 6' and 'David is 4'. The initial clinical diagnosis was that these answers were the result of treating an overlap problem as if it were a non-overlap one - it was assumed, in other words, that children failed /
failed to appreciate that one interval was included within the other.
A more general definition (but one which lacks explanatory power) is that
the overlap error is one in which the subject misconstrues the nature of
the relationships as given in the problem.

Donaldson initially classified both overlap error and asymmetry
error (to be defined later) as structural errors (that is to say, errors
which were the result of a failure to understand what was involved in the
problem or in part of it) to be distinguished from 'executive errors' or
slips.

The immediate jumping-off point for the present investigation was to
discover whether this classification of asymmetry error and overlap error
as structural errors held for simpler problem-forms. For if the aim is
to discover whether a given type of error is or is not structural, then it
would seem clear that the structure of relationships in which one is interested
must be presented in the simplest possible manner, with no added complexities
to obscure the issue. Then, if all attempts to present the bare bones of
the problem fail to prevent the occurrence of the error, this will support
the view that it is structural. However, it must be added that if the
error does disappear in these circumstances, it does not follow with equal
certainty that it is executive. Here, as so often in research, one result
may be decisive while another leaves the problem still open. A structural
error was defined (see above) as one which results from a failure to
understand what was involved in the problem. Now suppose that the subject
is given a simplified overlap problem followed by the more complex B2 (see
P23), that he has grasped the structure of the simplified problem-form, and
that /

---

\[\text{\footnotesize She also distinguishes a category which she calls 'arbitrary error' but,}
\text{perhaps because of the relative simplicity of the problem-forms used in}
\text{the present study, these were very infrequent indeed. They involve, she}
\text{says, "...a lack of loyalty to the given" (ibid, P202).}\]
that he has executed this problem without error; it may still be the case that when he comes to the more complex problem-form of B2 he may then fail to appreciate the structure of the problem, and his error may therefore be structural rather than executive.

The first attempt to simplify the original problems yielded a problem-form of which the following is an example:

We want to find out the ages of two girls called Jean and May. We know that a third girl, Betty, is 15 and that she is three years older than May and five years older than Jean. How old are Jean and May?

We call this problem B4.¹

If the reader will compare B4 with problems B1 and B2 he will notice that three major changes were made. The 'piece of information' requirement was omitted since it seemed to call for a different form of reasoning from that primarily involved in the overlap situation; the multiple-choice form was omitted (for reasons given by Donaldson (1963, Pp.109-110)); and the form of the problem was altered by specifying at once the terms involved in each of the relationships with the link (thus, I state, '...and that she is three years older than May', rather than '...and that she is three years older than one of the two girls, etc.').

Within this problem-form, systematic manipulation of the relevant variables was carried out. The two variables which seemed at that time chiefly relevant were the relative size of the first and second intervals, and the nature and direction of the relationships between the terms. These relationships were varied as before by using ages and heights; they were now further varied by stating the relationships in both possible ways (e.g. 'older than' and 'younger than'). Combination of these criteria generated eight versions of the problem:

¹ The system of numbering the problems here follows that adopted in Appendix I of 'A Study of Children's Thinking'.
TABLE 1

<table>
<thead>
<tr>
<th>Relative size of first interval</th>
<th>Small</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older</td>
<td>B4</td>
<td>B6</td>
</tr>
<tr>
<td>Younger</td>
<td>B8</td>
<td>B10</td>
</tr>
<tr>
<td>Taller</td>
<td>B11</td>
<td>B9</td>
</tr>
<tr>
<td>Smaller</td>
<td>B7</td>
<td>B5</td>
</tr>
</tbody>
</table>

Each of the relationships (e.g. 'older than' or 'younger than', etc.) remained constant within each question, that is 'older than' and 'younger than' were never used in the same problem. The sex of the names used in the problems was not considered crucial and hence was not varied exhaustively. Common pre-names were used and the sex of these was alternated throughout the series. Thus the names in B4 were female, in B5 male, in B6 female, and so on.

As a further example of the problems generated by this procedure, here is B7:

We want to find out the heights of two boys called George and Jack. We know that a third boy, Tom, is 4 feet 8 inches, and that Jack is 2 inches smaller than Tom and that George is 3 inches smaller than Tom. How tall are George and Jack?

These problems, together with B1 and B2, were given individually to twelve children whose ages ranged from 8 years 2 months to 11 years 6 months. We call these children Group A. Since not all problems were given to all children /
children, and since the sample was very small, no comparative counts of errors could be made; but I was able to examine the types and general trend of errors as they occurred throughout the sample.

Results - Group A

Overlap error was not eliminated by the simplification - indeed it occurred more frequently than any other type of error. Only one child, however, persisted in making the error in spite of attempts which were made to help the children overcome the difficulties. In this case, after overlap error had been made in the initial attempt, I covered one part of the question, and the child calculated the value of the dependent term correctly (in the part, 'We know that a third girl, Betty, is 15, and that she is three years older than May'); this part was then covered and the other relationship '

...and five years older than Jean' revealed, whereupon the child calculated Jean's age correctly. Finally I asked the child to do the whole question again, and she at once fell back into overlap error. She was quite unable to deal correctly with both relationships when they were presented within the context of a single question. In this case, then, overlap error would seem to be structural rather than executive in nature.

However, it occurred more commonly on the original than on the newer simplified problem-forms. Ann, for example, attempted the problems in the order B8, B9, B2, B11 and B10, and made no overlap error until she reached B2. The structure of B9 is (apart from the decrease in complexity) identical with that of B2 - both concern heights, both use the 'taller than' relationship, and in both the longer interval comes first. This happened also /
also in the case of David. Here overlap error on B2 was preceded by correct solution of B7, B8, and B9 in that order. The tendency to overlap error on the more complex form of the problem was thus particularly strong in both these cases. And since B9 and B2 do not differ in their formal structure, the error in B2 in these cases may have been of an executive kind. (In other cases where B2 was administered, it was part of an all-correct series.)

One child made asymmetry error on the first two problems (B4 and B6), corrected it on the third and fourth (B8 and B10) and returned to it on the fifth question (B1). It will be noticed that B8 and B10 are both 'younger than' problems, and B4, B6 and B1 all employ the 'older than' relationship. It is possible, therefore, that this pattern of asymmetry error is due to the type of relationship ('older than' or 'younger than') involved - an explanation which is strongly reinforced by the results reported later (see pages 81 ff.).

The trend of errors throughout the sample was puzzling. I confirmed the finding, reported by Donaldson, that overlap error tends to fall chiefly in the middle of the age distribution (but see results of Group B). In the present study, it was made by the children who were approaching ten, or who were ten, but not by the eight- and nine-year-olds nor by the eleven-year-olds. But unlike the findings reported in chapter 5 of Donaldson's book (where overlap error was usually preceded by some more primitive error) it was here found that with one exception the younger children achieved correct answers.

1 But need not have been: see discussion on page 24.

2 The nature of this error is fully discussed in chapter 4. Donaldson defined it as "...the treatment of an asymmetrical relation as if it were a symmetrical one". For example, if he had made an asymmetry error a child would give thirteen as his answer to the question: 'Jill is 10. Jill is three years older than Betty. How old is Betty?'
answers. This sample was so small and so unrepresentative that it was decided to replicate the experiment using larger numbers of children at each age level and using a group form of the test. These children formed Group B.

Group B

It is important to notice that, although the children of Group B were tested by group methods rather than by clinical methods, the purpose of this part of the investigation was purely exploratory. Now all that can be achieved by exploratory investigation, whether it employs clinical or group techniques, are hypotheses for later independent testing. If the group test findings happen to be striking or unexpected, it is more likely that useful hypotheses will result, but there obviously is no way of knowing whether this will be the case beforehand. The status of the group test at this level is thus identical with that of the clinical technique; the two can be considered complementary, the group test giving a broader view, and the clinical a more analytic one.

We have available two sets of group test results, one of which was the original check indicated above, and the other of which was a larger administration of an identical test given for another purpose. Since the two tests are exactly similar in content, and since the latter administration was to a larger and more representative sample, it is the results /

1 Unlike the group tests to be described later in which the aim was to test specific hypotheses derived from the clinical studies.
results of the latter investigation which are quoted in illustration.
(The results of the two investigations are substantially similar.)

On the occasion of the initial testing which we are considering here, the sample consisted of 890 children aged 7 years 11 months to 11 years 11 months inclusive. These children comprised the complete population of seven single-stream Primary Schools which (the Chief Education Officer claimed) were fairly representative of Preston primary school children.

One of the modifications of the test items suggested by the results of the study discussed above (Group A) was a simplification of the arithmetic involved in the solution. Two (and perhaps three) of the errors in this pilot study had been due to arithmetical difficulty, so it was decided to use no number higher than ten. As an example of the problem-form used, I quote Bl2:

We want to find out the ages of Jean and May. We know that Betty is 10, and that she is one year older than May, and three years older than Jean.

So Jean is ....... years old.

So May is ....... years old.1

A further modification thus included the dropping of 'two' (girls) and (we know that a) 'third' (girl). This was done because of the possibility which had by then been discovered in the unquantified three-term series problems that the children might achieve a correct solution of the problem by thinking in terms of four girls rather than of three. It was thought desirable /

---

1 The complete set of items is given in Appendix III
desirable to specify in advance as little of the mode of solution as possible so that the thought-processes in as 'pure' an overlap situation as possible could be studied.

In order to cut the length of the group test, it was decided to omit the 'heights' problems\textsuperscript{1}, but otherwise to leave all four variations\textsuperscript{2} as given in Table 1. These four problems are called Bl2, Bl3, Bl4 and Bl5.

Spacer problems\textsuperscript{3} of a related kind were interpolated with the overlap problems, and the eight items thus obtained were permuted in cyclical fashion to give eight forms of the test. These were then distributed in this order to the class as it sat.

The results of this group testing are represented in Figure 1. (Since statistical techniques are not appropriate in exploratory investigations, the results are presented in graphical form only.)

From the graphs it can be seen that the pass-fail-pass pattern from Group A was not confirmed. Although there was an increase with age in correct solutions, the slope was a very gradual one, almost linear in form, but with a slight negative acceleration. It will also be noticed that proportionally the incidence of overlap error is small - smaller than we might have expected as a result of the clinical studies (see also later discussion/)

\textsuperscript{1} Another reason was that the youngest group of children had not yet learned to deal in feet and inches.

\textsuperscript{2} If the reader will compare forms Bl2 and Bl3 in the Appendix, he will notice that we did not at this stage equate the number of times that the link was mentioned in different types of problem. In later forms, however, this omission was rectified - see page 37.

\textsuperscript{3} The spacers were of course intended to reduce interference (or facilitation) effects among the overlap items proper.
FIG. 1. RESULTS OF EXPLORATORY GROUP TEST

CURVE A: PERCENT CORRECT RESPONSES: Average (grouped in four-month intervals by age) over four problems.

CURVE B: PERCENT OVERLAP ERRORS: Average (grouped in four-month intervals by age) over four problems.
Another feature of these results is the uniformity in incidence of overlap error with age. The significance of this rather puzzling phenomenon was sought by further clinical study, and (in a later connection) a discussion of the issues involved is given on pages 118ff.

In an attempt to explain the several puzzling aspects of these results and at the same time to refine further the design of the overlap problem, I returned to clinical studies as a source of hypotheses. A random sample of 63 of the children was selected from the original cross-sectional study and these were interviewed individually by the usual 'thinking aloud' method.

**Results (individual testing): Group B**

In the individual interviews the child was presented with the problems - usually several weeks after the original group test - and asked to solve them again, but this time while thinking aloud. The complete interview was tape-recorded and later transcribed.

The first thing that was noticed was that - as in the group test - many of the younger children could solve the problems while many of the older children could not do so. We were, and to some extent still are puzzled /
puzzled by the successes of the younger children. It may be that their successes were spurious, based on a failure fully to understand what was involved in the problem-situation but the thinking-aloud procedure gave no direct evidence for this. Conversely, many of the failures of the older children appeared to be due to structural difficulty.

A good example of failure in older children is that of Jessie, a nine-year-old who was in the 'A' stream of a four-stream school. Jessie had made a classical overlap error in the following problem:

Bl4: ...We know that Peter is 8 and that he is three years older than Ian, and one year older than David.

The experimenter then asked:
E: 'Who will be one year older than David?'
S: Ian.
E: Ian, I see. It's definitely Ian, is it? (Pause) It couldn't be Peter, could it?
S: No.
E: Why couldn't it be Peter?
S: Because it says that ...see...that "...and that he is three years older than Ian and one year older than David."
E: So could it be Peter?
S: No.'

It /

1 Much of the present research is, of course, directed towards achieving a problem-form which avoids this kind of ambiguity.
2 The transcripts of the interviews conform to the conventional notation E = experimenter and S = subject. The interviewing technique which I used differed to some extent from that used by Donaldson. Whereas she attempted to find only the reason behind a child's spontaneous statement, I not only attempted this, but also tried to discover just how firmly embedded the error was in the child's thinking. I tried to shake the child out of his error by such means as attempting to show him contradictions in his reasoning and by diagrammatic representation of the problem. (Throughout this procedure the child had available a typed statement of the problem.)
It is fairly certain, then, that this older (and intelligent) child was experiencing a real structural difficulty with overlapping relationships.

Among the oldest group (ten-year-olds) there were many overlap errors also of an apparently structural kind. Here is an example from the 'A' stream:

**Janice - Problem B15:** We want to find out the ages of Bill and Harry. We know that Jim is 7, and that Bill is three years younger than Jim, and that Harry is two years younger than Jim. How old are Bill and Harry?

S: 'Well, Jim is seven, and Bill is three years younger than Jim; well, that means he is four.

E: Good.

S: And that Harry is two years younger than Jim. (Pause) That means he's two.

E: Mm. How do you know that?

S: Well you take the two from the four years that Jim is.

E: How old is Jim?

S: He is four.

E: How do you know that?

S: You take the three from the seven...

E: How old is Jim?

S: Seven.

E: Just do it once more for me, will you please? Start again.

"We /
"We know that Jim is 7, and that Bill is three years younger than Jim", so how old is Bill?

S: Four.

E: Uh-huh; and then Harry is two years younger than Jim, so...

S: That is two. You take the two from the four.

E: Now you said that Jim was seven, and then that he was four.

S: Now which is he?

E: Jim is seven.

S: Uh-huh.

E: And Bill is three years younger. Bill is four.

S: And Harry's two years younger than Jim, so how old is Harry?

E: Two.

One of the features of this protocol is the persistence with which Janice continued to make overlap error in spite of my prompting. Persistence was one of the chief criteria used in classifying overlap error in clinical interview (and other errors) as structural, and another example is given in this quotation from Graham’s protocol (Graham was aged ten):

B13: We want to find out the ages of Anne and Susan.

We know that Joan is 9 and that Susan is two years younger than Joan, and that Anne is three years younger than Joan. How old are Anne and Susan?

S: /
S: (Pause) 'Susan is seven. And Anne is - three.

E: How do you know that Susan is seven?

S: (Pause) "We know that Joan is 9 and that Susan is two...", so I took two from nine and that leaves seven.

E: Ah, good, yes. How did you find out Anne's age?

S: Then I took...then I took another three from seven and that leaves four.

E: Mmm, I see. So Anne is four, is she?

S: Yes.

E: I see. It says Anne is three years younger than Joan; how old is Joan?

S: (Pause) Nine.

E: Nine. So if Anne's three years younger than Joan, how old will Anne be?

S: Six.

E: Six. And if Susan is two years younger than Joan how old will Susan be?

S: Four. (That is overlap error again but with terms reversed, since we started with the other term first - an example of how, even if the child's attention is drawn to the fact that he should subtract from the link, he will continue to make overlap error.)

E: Uh-huh. How do you know?

S: (Long pause).

E: Why do you take two from six?

S: Because it's much easier.

E: Easier than what?

S: Taking six from two. (!)

E: But why from six? Why didn't you take two from nine? Susan is two years younger than Joan, isn't she? And Joan is what?

S: /
S: (Pause) Is nine.
E: Mmm. So Joan is nine and Susan is two years younger than Joan...
S: She is seven.
E: Seven. Mmm, that's fine. (Although Graham did not solve the overlap problem as a whole, I felt that to continue would disturb rapport.)

On the other hand, many of the errors among the older children were ambiguous, and were corrected by a very simple rephrasing of the question. Isabel, for example, made the classical overlap error by subtracting from the value of one of the dependent terms instead of from the link. When asked how she obtained her answer she made a clear statement to this effect. However, when the interviewer re-read the question for her fluently and with proper emphasis she was at once able to see that she should have returned to the link rather than to one of the dependent variables in order to calculate the value of the other dependent variable. Once she had achieved this insight she was quite certain that she had been wrong before, and that this was the correct way in which to tackle the problem.

In the course of work with these subjects, we became aware of the importance of the fact that in some of our problems the link was mentioned only once. So we now devised a third simplified form, in which the link was always repeated. We illustrate this form by Bl6:

Betty /

---

1 This child's error may seem at the beginning of the protocol to be executive, but by the end of the quotation it becomes clear that it is structural.

2 Hitherto the link in problems involving an 'older than' relationship was mentioned only once; but in problems involving a 'younger than' relationship it was mentioned twice (see page 31, footnote 2).
Betty is 10. Betty is three years older than May. Betty is five years older than Jean. How old are May and Jean?

(The other versions of this form are known as Bl7 - Bl9 inclusive.)

Here, if overlap error still occurs, it is less likely to be due to an executive error such as loss of hold of the link. The fact that it did still occur provides further evidence that the overlap error is of a structural kind. The problem was evolved during the clinical study which we have been discussing, and was given to the latter part of that group.

Donald tackled the problem which we have just quoted. Donald is aged eight:

S: 'Ten and three is thirteen. So May is thirteen.

E: What else?

S: Thirteen plus five is eighteen. So Jean is eighteen.

E: Yes. Now let me hear you doing the first part again.

S: Betty is ten. Betty is three years older than May, so May is thirteen. Betty is ten. Betty is five years older then Jean, so Jean is eighteen.'

Here then we have a very clear example of the conventional overlap error in spite of the fact that the child in his last statement has mentioned the link twice and has given it the same value on each occasion. The other error in this protocol is, of course, asymmetry, and I spent the rest of this particular session trying unsuccessfully to help the child achieve insight /
insight into asymmetrical structure. Donald's answer to his next problem, quoted below, will be given at length, since it gives us unusual insight into the mental processes involved in his solution. Here is the problem:

Bl7: Bill is 7. Dick is two years younger than Bill. Jim is three years younger than Bill. How old are Dick and Jim?

When Donald had made the classical overlap error I wrote out the question again in a slightly different form:

Bill is 7. Jim is three years younger than Bill.

Bill is 7. Dick is two years younger than Bill.

E: 'How old is Jim?
S: Jim is three years younger than Bill so Jim is four. Dick is two years younger than Bill. So Dick is five.
E: Now let's go back to this one (i.e. to the original question).
S: Bill is seven. Dick is two years younger than Bill, so Dick is five. And Jim is three years younger than Bill, so Jim is two.
E: How many boys were there here? (indicating the original form of the question).
S: Five.
E: Do it again and see (indicating special form of the question).
S: Bill is seven. Jim is three years younger than Bill, so Jim is four.
Bill is seven. Dick is two years younger than Bill so Dick is two.
(The/
(The experimenter managed to correct this by hiding one part of the question at a time.)

E: How many boys are there then?
S: Three. (The subject then made overlap error again on the original question.)

Later the same day I sent for Donald again and asked him to attempt the original form of the question once more. On this occasion he again made overlap error. On being asked why Jim was two, he replied:

S: 'I've got Dick's age. You take away three years from Dick's age which is two.
E: How do you know to take three years from Dick's age?
S: Because he is nearest youngest to him.
E: Why do you take it from the nearest youngest?
S: You can't take it from the biggest because you'll get the wrong answer.
E: How is that?
S: Because you'll get four and that wouldn't be right, so you take it away from the other age.
E: Yes, but why?
S: Because it's the wrong answer and I done it before and got it right then.
E: Could four be the right answer?
S: Yes if there was only two boys. But there's three boys.
E: Which two boys do you mean?
S: Bill and Jim.
E: Why do you think it's wrong with three boys? (No reply).
S: /
S: Bill is five so Jim is two.

E: You say Bill is five. It says Bill is seven here. Have you got two Bills with different ages?

S: Yes.

E: If both Bills were the same age - if they were both seven - what would be the ages of Dick and Jim then?

S: Dick would be five and Jim would be four.

E: If both Bills were the same boy, what would happen then?

S: Dick would be five and Jim would be two.

There are many interesting aspects to this protocol, among them the explicit statement that Jim could be four if there were only one pair of terms. But the most interesting insight of all is that the correct solution can be achieved (at the end) only if there are two Bills of the same age. As soon as the two Bills become one, the overlap error appears again. This supports very strongly the hypothesis that the difficulty is in returning to the link.

While overlap error on a simplified problem of this type is almost certainly significant for our purposes, a correct solution on the other hand is ambiguous. Just as in the earlier problem-form we suggested that the younger children might be achieving spurious success by separating the two premisses, so a fortiori in the case of these simplified problems. In these /

---

1 Here of course the child may just be agreeing with the experimenter.

2 See discussion on pages 157 ff.
these problems it is even easier to achieve a solution without conceiving that the relationships overlap (or that one is returning to the link) since the link is mentioned twice and the two sets of relationships can thus be dealt with separately.

To prevent this happening we attempted to construct problem-forms in which the subject, to achieve solution, would be forced to consider the overlapping of the relationships. We have called these problems 'partially quantified' because the link (Betty in B20 below) is given no absolute value. Here, as an example, is B20:\n
Betty is 3 years older than May. Betty is 2 years older than Jean. What is the difference between the ages of Jean and May?

The trouble with this problem is that the difficulty of dealing with the partially quantified form has in practice occasionally overshadowed the overlap difficulty. Sarah, aged nine, for example, interpreted the statement, 'Robert is one year younger than Sam', as 'Sam is one year younger than Robert' - the usual asymmetry error. The experimenter then asked:

E: /

---

1 This might account for the lack of discrimination in fully quantified problems, although it does not explain the direction of the significant partially quantified findings (see analysis of variance on pages 68 and 69).

2 The usual variations of B20 are known as B21, B22 and B23.
E: 'What does it say in the second line?
S: Robert is one year younger than Sam.
E: Does that mean Sam's one year younger than Robert?
S: No.
E: What does it mean? (Pause) If you compare Sam with Robert, what do
you say? Sam is what than Robert?
S: Older.
E: How much older? (Pause) Can you tell me how much older Sam is than
Robert? (Pause) Any ideas?
S: No.
E: How much younger is Robert than Sam?
S: One.
E: One year. Robert's one year younger than Sam, is he?
S: (Pause) Yes. (I then wrote down the unquantified statement, 'Robert
is younger than Sam', and asked what could be said about Sam. Sarah
supplied, "Sam is older than Robert", and this was written down. I
then wrote, 'Robert is one year younger than Sam', and asked what followed)
E: How much older is Sam than Robert? Sam is so much older than Robert.
How much is it going to be?
S: (Long pause) Five.
E: You mean that Sam is five years older than Robert? Is that what you mean?
S: Yes. (She appeared to be unable to make any more progress without help,
and so I suggested an absolute age for one of the variables:)
E: Let's have a look at it this way. Just for fun we'll say that Robert is
ten. What age would Sam be?
S: /
S: Eleven.

E: Eleven—good for you. Sam is eleven, right?

S: Yes.

E: Now if Sam is eleven, how much older is Sam than Robert? (Pause) If Sam's eleven and Robert's ten, Sam is how many years older than Robert?

S: One.

E: Uh-huh, that's right. 'So Sam is one year older than Robert'. Do you understand that?

S: Yes. (After this, I went back to a partially quantified relationship and Sarah once again failed to reverse it. I therefore gave another illustration using the fully quantified form. The question then became, 'Robert is six. Robert is three years younger than Sam'.)

E: How old will Sam be?

S: Nine.

E: So how many years is Sam older than Robert?

S: Three.

E: Now if we make that, 'Robert is four years younger than Sam', how many years is Sam going to be older than Robert?

S: Four.

E: Mm. Why four?

S: Em, Robert is four years....

E: Uh-huh? Younger?

S: Younger than Sam. So Sam would be eight.' (It can be seen that Sarah, at the beginning, was quite unable to deal with partially quantified premisses. Towards the end of the interview, she was beginning to understand what was involved.)
The fact that these difficulties (and they are not uncommon) occur when dealing with a single asymmetrical relationship indicates that the difficulty is specific to the partially quantified situation, and has nothing to do with overlap difficulty per se. For this reason it is probably safer at this stage (and it may always be necessary) to use the partially quantified form along with the simplified fully quantified form. The two problem-forms together contribute the information which either alone is unable to provide.

It soon became obvious, however, that there were loopholes in this (partially quantified) form too. First of all children tended to solve it by attributing an absolute age to the link and then solving it as if it were a fully quantified problem. The children who did this fell into two groups—those who treated the adoption of an absolute age for the link explicitly as a conditional procedure, and those who believed that they were able legitimately to discover the value of the link from the information which is given. An example of the former kind of solution follows:

Isabel, who was presented with the following problem:

Jill is 3 years older than Sally. Jill is 1 year older than Rose.

showed that she was aware of the status of the value she attributed to the link:

S: 'If Jill was ten, Rose would be nine and Sally would be six.

E: Mm. How do you know?

S: Because if Jill was really ten - if she was ten or not - it would be right, because Sally would be youngest and Rose would be second oldest.

E:
E: (Later in the same protocol:) Well, could you work it out again for me and see if you can do it?

S: With a different...

E: No. We'll still say Jill is ten.'

An example of the latter kind of solution was given by Charles, who was presented with the same problem as Isabel. He began by saying that Jill was three:

S: Oh, Jill's three.

E: How do you know Jill's three?

S: It tells you there.

E: Read out that bit.

S: "Jill's three years older than...S...than Sally". (Note his distress)

E: Does that mean Jill's three?

S: No.'

This answer, which might superficially seem to be merely a mis-reading, is so common that it seems to spring from an anticipation of a fully quantified situation, rather like the anticipation of a non-overlapping structure suggested by Donaldson. Joanna, in her attempt to solve a partially quantified problem whose premises are stated in the opposite direction from those of the problem we have just discussed made the same kind of error:

Tom is 4 years younger than Dick. Paul is 3 years younger than Dick. What is the difference between the ages of Tom and Paul?

S: /
S: 'Well, Tom is four years younger than Dick and Paul is three years younger than Dick.
E: What is the difference between Tom and Paul?
S: (Pause) One year.
E: How do you know that?
S: 'Cos you know that Tom is four and Paul is three and take three away from four.
E: How do you know that Tom is four?
S: It says that Tom is four years younger than Dick.
E: Does that mean Tom is four?
S: No.'

Thus, by converting the relatives to absolutes, she obtained the correct answer without tackling any form of overlap problem - either fully quantified or partially quantified.

Jessie made exactly the same error but failed to see her mistake when a direct question was put to her:

S: 'Tom is four and Paul is three and Tom is one year older than Paul.
E: Now are you sure that Tom is four? Do you think that Tom is four?
S: Yes.
E: How do you know that Tom is four?
S: (Pause) Because it says that Tom is four years younger than Dick.'

Since the child went on working on this assumption it seems that, at least in this case, the error was the result of very much more than a mere misreading.

Children /
Children who converted partially quantified problems into fully quantified problems in this way were often asked whether a change in the value of the link would lead to a change in their answer - that is, to a change in the difference between the two dependent variables. Many children believed that the difference would in fact change with a change in the value of the link term. Here, for example, is an excerpt from the protocol of Francis:

S: 'Well, if Jill is ten then if Ann...Rose is one year younger, she must be nine; and if em... and Jill is three years older than Sally, Sally must be seven, and the difference between Rose and Sally is em... two years.

E: Lovely. Now if Jill is eight would there still be two years between them?
S: No.
E: Why not?
S: Because then ... Rose would be seven ... and Sally would be five, and then the difference would be ... two years!'
E: It's the same again, isn't it?
S: Yes.
E: Now if Jill is nine the difference would still be two years?
S: Yes.
E: Why is it always two years? That's a puzzle, isn't it?
S: It's because ... (Long pause)
E: Have you any ideas?
S: (Pause)
E: Why is it always two years?
S: (Long pause) I don't think I do know why it is always two years.'
Less obvious - but more serious perhaps - is the fact that in this form of the problem children use the word 'difference' as a signal to subtract, without really understanding why they should subtract. Thus, in the questions as they stand, they are able in this way to achieve the correct answer without comprehension. Here, as an illustration, is an excerpt from John's protocol:

Jill is 3 years older than Sally. Jill is 1 year older than Rose. What is the difference between the ages of Sally and Rose?

S: 'Two.
E: Good. Now tell me how you worked it out.
S: If it's 'difference' you always have to take away.
E: Oh, I see. This word here, 'difference': is that why you took away?
S: Yes.'

The fact that John did the same thing in a non-overlap situation (and thus obtained the wrong answer) reinforces our interpretation of his statement:

Ruth is 3 years older than June. June is 2 years older than Kate. What is the difference between the ages of Ruth and Kate?

E: 'Now let me see if you can do it. (Pause)
S: One.
E: How do you know it's one?
S: Two from three leaves one.
E: Why did you subtract?
S: To find the difference. To find the difference you have to subtract.

Euphemia /
Euphemia made the same error in a non-overlapping three-term series, and explained it as follows: 'Because the difference means take away and I have taken away'.

The problem-form was therefore modified again (generating versions B24 - B27 inclusive) and rephrased as follows:

**B24:** Jill is 3 years older than Sally. Jill is 1 year older than Rose.

Who is older, Sally or Rose? Write your answer here ......

How many years are there between Sally and Rose? Write your answer here ......

This form, it will be noticed, avoids the difficulty inherent in the previous form by omitting the word 'difference', but there is some evidence which suggests that this form is still not perfect.

In his attempt to solve a non-overlapping three-term series problem, Thomas subtracted and therefore obtained the wrong answer. He went on:

S: 'Please sir, I was thinking, adding it, but it wouldn't come out right.

E: Why wouldn't it come out right if you added it?

S: Because it says how many years are there between Ian and Jim, and if you added them that would be putting them together, instead of splitting them up.' (I asked Thomas why 'between' meant splitting them up, but he was unable to answer this.)

A solution to this difficulty was discovered later (see page 54).
One of the crucial tests of overlap difficulty is to compare the performances of matched groups of children on overlap and non-overlap problems. Before this was feasible, however, it was first necessary to study non-overlap problems clinically, both in an attempt to discover the mental processes involved and in order to match accurately the overlap and non-overlap problem-forms to be used in the group test.

The clinical analysis of non-overlap problems foreahadowed to some extent the surprising result which the group test was to produce: in their solutions to these non-overlapping problems some children produced solutions which involved overlapping structures. In other words, in the same way that children who attempt overlap problems sometimes produce a solution including a non-overlapping structure (or 'overlap error') so children attempting non-overlap problems sometimes did the reverse. Our important issue then is whether the overlap error produced in the non-overlap problems is of the same (structural) kind as that produced in the overlap problems; or whether it is the result of some spurious characteristic of the items used. The reader will recall that in the case of the overlap problems, a long process of simplification was undertaken to ensure that the overlap error was not due to anything other than the intrinsic nature of overlapping relationships. A similar (but, as the result of our knowledge in the previous stage - telescoped) study has been made in the case of non-overlap problems. The initial clinical study utilised four non-overlap problems, two of which were partially quantified and two fully quantified; within each type, one problem used 'older than' relationships, and the other 'younger than':

\[ \text{TABLE 2} \]
As an example of the fully quantified form here is Question 1:

Betty is 8.
Betty is three years older than May.
May is one year older than Jill.

What are the ages of May and Jill?

And as an example of the partially quantified version here is Question 2:

Jean is 3 years older than Sally.
Sally is 1 year older than Rose.

Who is older, Jean or Rose?

How many years are there between Jean and Rose?

The phrasing of the latter question was shown to be misleading and to lead to spurious overlap error. In his answer to the first part of Question 2, Sinclair said:

S: /

---

1 In this version, of course, the problem invites asymmetry error. An alternative version of Question 1 was also produced which avoided asymmetry difficulties. This was given to those who experienced asymmetry difficulty in their initial interview (Question 3 was in the non-asymmetry version).
S: 'Jean.
E: Why?
S: Jean is three years older; Rose is one year older (note his asymmetrical interpretation of the second relation).

Two years between.
E: Why?
S: Three take away one leaves two.
E: Why do you subtract?
S: 'Between' means take away.
E: Are you sure?
S: Yes.'

For this reason we decided to adopt a different form for the second question. Instead of 'How many years are there between Jean and Rose?' we substituted 'How much older is she?' This version might a priori appear ambiguous, but in practice, both clinically and experimentally, it appeared to cause no difficulty; and it has the great advantage of being applicable no matter which alternative the children select in the first question. (The problem-form remained otherwise unaltered.)

However, the non-overlap problems continued to provide instances of overlap error in spite of this modification.

In her answer to Question 2, version 2, for example, Alison said:

S: 'Jean.
E: Why?
S: She's older than Sally and Sally is older than Rose (note the sophistication /
sophistication of the seriation). Two years older — I took the one off the three and it made two.

Although no other reasons were obtained in such instances, it seemed to the investigator that the lay-out of the question might possibly have an effect on the child's method of solution: in each case the problem was set out with one statement directly under the other; in every instance, the larger value was on top; and in the case of the partially quantified problems the values were given as numbers. Now while there is no logical reason why this should lead to subtraction, it could be the case that psychologically this predisposed the children to subtract. To check this possibility, two new forms of problem were evolved: the first kind was used only with those children who had subtracted in the first instance, and purely as a check on this hypothesis; the second kind was used as a permanent modification of the problem-form. In the former instance then the investigator simply presented the same problems on a second occasion but with the small number on top. No other modification was made in this instance.

Alison, who displayed a clear-cut, if erroneous, approach in the protocol described above, was given the new version a few days later. There is no direct evidence of course that her confused answer on the second occasion is due to the lack of opportunity for automatic subtraction, but this seems to be the most likely explanation.

The

---

1 This is not to say, of course, that the problem lay-out is the only cause of overlap error in non-overlap problems, as we shall see when the alternative form is used.
The 'check-version' was as follows:

Jean is 1 year older than Sally.
Sally is 3 years older than Rose.

Who is older, Jean or Rose?
How much older is she?

To the first question, Alison answered:

'Rose.

E: Why?
S: There is one year between Jean and Sally, and three between Sally and Rose.
E: So why is Rose older?
S: ... If Jean was about 5, Sally would be 6. And because Sally is 3 years older than Rose. If Sally is 8, I mean 6, Rose is 10. So Rose is 3 years older than Sally.
E: What's the answer to the second question?
S: Rose is 10.
E: Is that the answer?
S: Yes!

There is a marked difference in the quality of the two sets of answers, and this difference is reflected also in her answers on the two occasions to Question 4 which was treated in a similar way.

Two other children who had subtracted on the earlier occasion now added without apparent confusion. While this could be a practice effect, the evidence /
evidence as a whole suggests that a permanent modification of the test item, which nevertheless does not commit us to a statement of the smaller interval first, is worth while. The second and permanent new form conforms to these criteria. In this third version of the non-overlap problem values were stated in words instead of figures\(^1\), and the relational statements used no longer printed in two quite separate lines. Question 2, for example, became:

Jean is one year older than Sally. Sally is three years older than Rose.

Who is older, Jean or Rose?

How much older is she?\(^2\)

These new versions were then given clinically to see whether overlap error still occurred.

Overlap error did in fact occur, but (due perhaps to the relatively small number of children who tackled version 3 clinically) the reasons given for it in interviews were reasons which could apply equally to the overlap form\(^3\).

And /

\(^1\) with the exception of the link term in fully quantified problems.

\(^2\) In the new version of Question 4 the large interval was stated first for comparison purposes. While we did not have a complete permutation of possibilities, this was not considered crucial at the clinical level.

\(^3\) One possible reason for error in partially quantified non-overlap problems (and for that matter in overlap problems) was the occurrence of the processes which Hunter (1957) called conversion and reordering. In conversion, the subject converts the relation 'Bill is two years older than Ian' to 'Ian is two years younger than Bill' and in reordering, the subject changes the order of two relations, e.g. 'Bill is two years older than Ian' 'Ian is one year younger than Dick' becomes, 'Ian is one year younger than Dick. Bill is two years older than Ian'. A clinical study of the forms used in the present investigation revealed no need on the part of the children to use either of these processes in either overlap or non-overlap forms. One or two children did in fact occasionally use one of the processes in his explanation, but there was no consistency, and the majority avoided their use altogether. The problem forms themselves make the use of these processes unlikely; the premises in the partially quantified problems are isotropic and so neither reordering nor conversion should be necessary.
And if this were the case we would expect overlap error to occur equally often in the overlap form.  

Rosalind, for example, made several interesting overlap errors, but in each case these seemed to be due (in ex post facto explanation) to other factors than the relationships in which we are interested. Having subtracted in Question 2, for example, she says in explanation:

S: 'There's two years between them.
E: Why take away?
S: You must take away to find the difference in their ages.' (c.f. all our attempts to avoid using the term 'difference' in the statement of the question!)

Later that same day, the investigator called for Rosalind again and asked her to repeat Question 2. She again seriated the variables efficiently in justification for her (correct) answer to the first question. She continued:

S: 'Jean is four years older.
E: Why?
S: Jean is one year older than Sally, and Sally is three years older than Rose, and three and one is four.
E: Why did you add?
S: If you subtract you get the wrong answer.
E: Why would it be wrong?
S: /

---

1 As will be seen later in the group tests, overlap error was in fact significantly more common in the non-overlap version, as were errors of all kinds combined. There must therefore be other and intrinsic reasons for overlap error in non-overlap problems.
S: Because if you subtract you get ... (Pause) ... you will find out how old Sally was ... I think, anyway. And if you add you get the answer Jean is.

E: How do you mean?

S: The amount of years older she is.

E: And if you subtract?

S: You find out how much older Sally is.*

Although Rosalind gets the correct answer here, therefore, it is obviously not based upon a very sound grasp of the problem-structure.

In her repeat of Question 4, Rosalind demonstrated some interesting interpretations:

S: 'Donald is the youngest, because Donald is five years younger than Ian and Peter is only three years younger than Ian. So he is two years younger (than Peter, presumably). (Note that she treats the statement 'Ian is three years younger than Peter' as if it stated, 'Peter is three years younger than Ian', and hence ends up with an overlap problem. As stated earlier however - see footnote 1, page 58 - if this were the chief explanation of the overlap error, it should occur equally in the opposite direction in a group test and it does not.)

(I then asked her to re-read the question, whereupon she saw her error, and seriated correctly. She then proceeded as follows:)

S: 'How much younger is he? ... Two years. Because he's five years younger than Ian and Ian is three years younger than Peter. So take three from five and two is left.
E: Why subtract?

S: To find younger, not older.

E: Why do you subtract to find the younger?

S: Well ... If you are finding something that's older you add. To find something that's younger you've got to subtract since it's a smaller amount." (Yet note that she had subtracted on her initial attempt at Question 2, an 'older than' problem). (Again these considerations apply equally to overlap problems, so in the light of subsequent group test results they cannot be claimed as fully explanatory of the excess overlap error in the non-overlap form.)

Although Rosalind's errors were also made by other children, no other child made any new errors.
CHAPTER 3

Overlap Error: Experimental Design and Group Test Results

From knowledge of problem-statements in the overlap form and from the information gained in the clinical study just described, it seemed that the problem-form which had been evolved for non-overlap problems was sufficiently unambiguous to be used in a group test. A design was therefore evolved the aim of which was to discover the relative structural difficulty of overlap and non-overlap items.

The problems used in the group test differed from the forms described so far in one respect. The 'younger than' statements were followed by the question, 'who is older, X or Y', not, 'who is younger?'. The effects of this change had already been checked clinically, and they seemed relatively slight.

1 This is not to say that if we found non-overlap problems in fact more difficult than overlap, we would know unequivocally why this was the case; rather that we needed (and felt that we now were in a position to obtain by group test) information on the relative structural difficulty of overlap and non-overlap problems: The clinical experience with non-overlap problems had made us less confident about the outcome of an overlap - non-overlap experiment, but we still had no clear evidence on the issue even of a clinical kind.

2 Another respect was in their content; this was done to add variety and to try to avoid any set between problems in their methods of solution. While it would be interesting to examine the effects of this on their problem-solution (c.f. Piaget's décalages) the identical content was used in both overlap and non-overlap forms of the test, and barring any unexpected interaction effect we can be reasonably sure that it did not affect the results of the test.
slight. (This alteration had been introduced because change in content had made phrasing the question, 'Which is less, etc.', rather difficult):

In her answer to the problem:

Bill is five years younger than John. Dick is three years younger than John.

Who is older, Bill or Dick?

How much older is he?

Carol said:

'Bill is oldest by 2 years. Oh no! 'Younger'. Dick's older by 2 years.... 'Older' in the question made me think it was 'older' here.'

This was the only clear example, and as can be seen, the subject corrected her error spontaneously. Even if spontaneous correction does not occur, however, the only part of the problem to be affected is the first question; the last (and important) question is unaffected\(^1\). Again this form is used in both overlap and non-overlap versions.

The first step was to construct unambiguous problem-forms of each kind (overlap and non-overlap) which were matched for content and irrelevant structure. For reasons given earlier both fully- and partially-quantified forms were used and within each of these forms, both 'older than' and 'younger than' problems were constructed. The other variable included was relative size of the first interval (i.e. whether it was larger than or smaller /

\(^1\) Even if this is not the case, this question should bias the results in such a way that the overlap problems would turn out to be structurally more difficult. In fact, as will be seen below, the reverse is the case.
smaller than the second). These variables were combined in 16 overlap and 16 matched non-overlap problems in the following way:

**Fully-quantified/partially-quantified:** it was decided to give both kinds to each subject, but to present them in blocks: that is to say, each subject tackled all problems of one kind before tackling problems of another kind. (Since the problems were to be presented in cyclical fashion, this permutation was carried out within each kind of problem.) Successive children were presented with the fully quantified block first or the partially quantified block first alternately. These decisions were made since interleaving of fully quantified and partially quantified problems within tests might have robbed the partially quantified problems of their effects. (That is to say, children would have been more likely to treat partially quantified problems as fully quantified throughout, for example by attributing absolute values to the link; and, conversely, if partially quantified problems were always given before fully quantified problems, the partially quantified problems might have had some unknown effect on the fully quantified ones.

**'Older than'/ 'Younger than':** it was decided to 'interleave' these variables within subjects, and thus to eliminate their effects as far as possible. ('Older than' and 'younger than' never appeared together in the same question.)

**Relative size of first interval:** it was decided to vary this effect in alternate problems by pairs: that is to say, two problems with the larger interval first were followed by two problems with the smaller interval first, and so on. In this way coincidence of the 'older than'/ 'younger than' effect and the present variable was avoided.

**Asymmetry**
Asymmetry difficulties were avoided since the design was already complex, and since a separate experiment devoted to asymmetry error was to be undertaken (see pages 81 ff.)¹.

The sex of the names used in problem statement was alternated. (This was not considered an important variable.)

As already stated the test was rotated within two separate blocks, one of fully quantified problems and the other of partially quantified problems.

An example of one of these rotations follows²:

QA: Overlap/ older than³/large interval first/fully quantified
QB: " /younger than³/ " " "
QC: " / older than/ small interval first/ "
QD: " /younger than/ " " "
QE = QH as for QA - QD, but with different content
QJ: Overlap/ older than/ large interval first/partially quantified
QK: " /younger than/ " " / "
QL: " / older than/ small interval first/ "
QM: " /younger than/ " " / "
QN = QQ as for QJ-QM, but with different content.

The non-overlap form is identical with the overlap form in these and in all other respects.

In both cases, identical 'spacer' problems were interleaved with all questions, giving 32 items in all.

Here, /

¹ But see the subsequent design, incorporating asymmetry difficulties described on pages 149 ff.
² c.f. Table 1, page 26 in which a similar set of manipulations is summarised.
³ 'Older than' and 'younger than' are used archetypically for the appropriate expression under other conditions. Where appropriate read 'more than'/'less than' or 'taller than'/smaller than'.
Here, as an example, is Question D (overlap version): ¹

Joyce jumps 6 feet.  
Ann jumps one foot less than Joyce. Nancy jumps three feet less than Joyce.  
So, Ann jumps ______ feet.  
So Nancy jumps ______ feet.

These sets of problems were thus intended to vary systematically or equate those factors which were considered irrelevant, so that the main purpose of the test could be fulfilled, namely the comparison of children's responses to overlap and non-overlap problems in such a way that the results would not be specific to any of the above conditions.

ADMINISTRATION

The problems were given to matched groups of children whose ages ranged from 8+ to 10+. Twenty-six² children were selected at random³ from the complete age-group concerned. The test booklets were then distributed to the children as they sat, alternate kinds (overlap and non-overlap) and successive rotations being given to the children in succession.

Children /

¹ The complete set of problems is given in Appendix III, pages xv ff.
² The maximum number available in the smallest class.
³ Each child in the age-group was given a number, boys and girls separately; a table of random numbers was entered and children eliminated in this way from the group (boys then girls) until the required number was reached including as nearly as possible an equal number of boys and girls. The school had only one stream until the 10-year-old group was reached. Here proportional representation was made and children eliminated from each stream separately.

There seems no reason to suppose that these children are unrepresentative of the general population of primary school children in this country with regard to the variables in question.
Children were instructed not to go back to problems which they had been unable to tackle, but to attempt each problem ('puzzle') as they came to it. A comprehensive set of instructions was drawn up and used on each occasion of testing. Questions were allowed before testing began but not thereafter. No time limit was imposed.

FINDINGS

Several analyses were carried out on the results. The first of these was made in an attempt to discover whether there was in fact a significant difference in the numbers of overlap errors\(^1\) occurring in each treatment (i.e. in each problem-form).

By inspection it was decided that the incidence of overlap error in each of the fully quantified forms was too small to make comparison valid:

\[\text{TABLE 3/}\]

---

\(^1\) By 'overlap error' is meant misinterpretation of the inter-relationships of the statements as given in the problem statement. Overlap error, defined in this way, thus applies to both overlap and non-overlap three-term series problems. In the former, overlap error refers to the structuring of an overlap problem as a non-overlap problem; in the latter, to the structuring of a non-overlap problem as an overlap problem.

It is important to notice that the actual incidence of overlap error can only be inferred from the numerical solution (see discussion on page 86). Nevertheless, it is proposed to substitute for the clumsy phrase 'wrong solution presumably as a result of overlap error', the phrase 'overlap error' when referring to the numerical solution. Since much time was devoted to making the problem-forms unambiguous in this sense, I feel that I am justified in replacing the longer phrase by the shorter. Where a clear distinction between process and product is necessary, it will be made, by distinguishing the numerical solution from the error which presumably produced it.
TABLE 3

Overlap error in fully quantified problems

<table>
<thead>
<tr>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td>$\pm x = 1$</td>
<td>$\pm x = 0$</td>
</tr>
<tr>
<td></td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
</tr>
<tr>
<td>non-overlap form</td>
<td>$\pm x = 1$</td>
<td>$\pm x = 4$</td>
</tr>
<tr>
<td></td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
</tr>
</tbody>
</table>

It will be recalled that each subject attempted eight fully quantified problems. The possible value for any cell is thus 104. It seems likely, therefore, that the higher incidence of overlap error (with overlap problems) in clinical studies is due to the ‘thinking aloud’ situation - hence the need for statistical verification of clinically derived hypotheses. (The difference in incidence of overlap error between fully quantified and partially quantified problems can probably be explained by the fact that it is relatively easy to achieve correct solution to fully quantified problems without attempting to construct a three-term series at all - see pages 41f.)

The results in partially quantified problems follow. These are represented as curves C and D in Figure 2:
Fig. 2.  PARTIALLY QUANTIFIED PROBLEMS

CURVE A = OVERLAP GROUP - CORRECT RESPONSES
CURVE B = NON-OVERLAP GROUP - CORRECT RESPONSES
CURVE C = OVERLAP GROUP - OVERLAP ERRORS
CURVE D = NON-OVERLAP GROUP - OVERLAP ERRORS
TABLE 4(a)

Overlap error in partially quantified problems

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td>$\leq x = 0 \ (n = 13)$</td>
<td>$\leq x = 2 \ (n = 13)$</td>
<td>$\leq x = 2 \ (n = 13)$</td>
</tr>
<tr>
<td>non-overlap form</td>
<td>$\leq x = 20 \ (n = 13)$</td>
<td>$\leq x = 11 \ (n = 13)$</td>
<td>$\leq x = 20 \ (n = 13)$</td>
</tr>
</tbody>
</table>

$\leq x_{ov.} = 4$

$\leq x_{non-ov.} = 51$

$\leq x_{8} = 20$

$\leq x_{9} = 13$

$\leq x_{10} = 22$

An analysis of variance carried out on the figures from which these results were derived shows that the difference in the incidence of overlap error in the two problem-forms is highly significant.

TABLE 4(b)

1 See also discussion and results in Appendix I, pages 152 ff.

2 Since at this stage the partially quantified numerical results (and hence whether the intervals had been overlapped or not) were considered chiefly relevant, only the scores from the second question following the problem-statement were considered. The scores for the first question - 'Who is older?', etc. - were considered later.

3 Details of the calculation will be found in Appendix II, page 1.
From the table it can be seen that, as with asymmetry error (see page 90 and discussion on pages 118 ff.), there is no significant change in the incidence of overlap error with age, nor is there any significant interaction factor.

The direction of the results (that non-overlap problems produce significantly more overlap errors than overlap problems) is puzzling. From our clinical work on overlap problems, we discovered that children tended to structure overlap problems as non-overlap ones; as a result we assumed, wrongly as it happened, that non-overlap problems would be easier than overlap problems in the sense that children's responses to them would produce fewer (or no) overlap\(^1\) or other structural errors. This assumption should have been checked earlier.

When /

\(^1\) as defined on page 66.
When non-overlap problems were given clinically, I was able, as a result of the clinical experience gained with overlap problems, to telescope the procedure of evolving a satisfactory problem-form, and so tested relatively few children. As a result it is impossible at the moment to explain this finding; more clinical work with non-overlap problems will be necessary. (Both a priori and from the few instances in which children did make overlap error in non-overlap problems during clinical study, it seems that it will be impossible to specify further the essential characteristics of overlap error in non-overlap problems at the behavioural level — that is to say, that it will be impossible to specify it, as we had hoped to specify overlap error in overlap problems in terms of characteristics such as 'return to link difficulty' and 'inclusion difficulty' — see Appendix I, pages 157ff. Once the essential characteristics of an error have been specified, it becomes possible, in theory at least, to 'explain' the error in terms of a coherent body of theory — for example, Piaget's attempted description/explanation in terms of a system of logic; or in neurophysiological rather than in psychological terms.)

It seems then that only partially quantified problems are useful in distinguishing the relative incidence of overlap error in children's responses to the two forms of three-term series problem.

Two other kinds of (presumably) structural error occurred in children's group-test solutions to three-term series problems. The first of these is failure to respond at all. Because of the nature of the clinical /

---

1 This has since been carried out but the protocols give no further indication of the reasons for this difference. It appears that clinical study (perhaps for the reasons given earlier — see Chapter 1) does not always give the information required (see also discussion on Pp. 117 ff.)
clinical situation, we have no clinical evidence on the reasons for this behaviour, but \textit{a priori} it seems reasonable to assume that children fail to respond to these problems because they fail to understand the structure of the problem: in other words, the error should be classified as structural rather than executive. (It seems to be the case that children were adequately motivated, since in every case children who failed to respond at all to the three-term series problems had completed the spacer problems.)

The frequencies of failure-to-respond on each problem-form are given in the following table:

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|}
\hline
\multicolumn{1}{|l|}{\textbf{Failure-to-respond in partially quantified problems}} & \textbf{age 8} & \textbf{age 9} & \textbf{age 10} \\
\hline
\textbf{overlap form} & \(\bar{x} = 34\) & \(\bar{x} = 27\) & \(\bar{x} = 1\) \\
& (n = 13) & (n = 13) & (n = 13) \\
\hline
\textbf{non-overlap form} & \(\bar{x} = 50\) & \(\bar{x} = 40\) & \(\bar{x} = 2\) \\
& (n = 13) & (n = 13) & (n = 13) \\
\hline
\end{tabular}
\end{table}

The /
The results are in the same direction as the results for overlap error. An analysis of variance\(^1\), however, shows that this difference is not a significant one:

\[ \text{TABLE 5(b)} \]

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form V</td>
<td>1</td>
<td>11.538</td>
<td>11.538</td>
<td>1.963(1,72)</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>non-overlap form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>2</td>
<td>140.334</td>
<td>70.167</td>
<td>11.937(2,72)</td>
<td>.01</td>
</tr>
<tr>
<td>Forms X Ages (Interaction term)</td>
<td>2</td>
<td>4.847</td>
<td>2.424</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within cells</td>
<td>72</td>
<td>423.230</td>
<td>5.878</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>77</td>
<td><strong>579.949</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(As one would expect, there is a significant decrease in the numbers of failures-to-respond with age; there is no significant interaction term.)

A number of children used one of the interval values as their numerical answer. Although this did not occur in the clinical study, it seems to be related to the treatment of relatives as absolutes (see pages 47 f).

Again it seems fairly reasonable to classify the present error as structural, although the issue of classification is not an important one for the present analysis. The frequencies of this error under each treatment /

\(^1\) Details of the calculation will be found in Appendix II, page ii.
treatment are given below:

**TABLE 6(a)**

Use of interval-value as answer in partially quantified problems

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td>$\bar{x} = 22$</td>
<td>$\bar{x} = 18$</td>
<td>$\bar{x} = 5$</td>
</tr>
<tr>
<td></td>
<td>($n = 13$)</td>
<td>($n = 13$)</td>
<td>($n = 13$)</td>
</tr>
</tbody>
</table>

$\sum_{x} = 45$

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-overlap form</td>
<td>$\bar{x} = 26$</td>
<td>$\bar{x} = 25$</td>
<td>$\bar{x} = 31$</td>
</tr>
<tr>
<td></td>
<td>($n = 13$)</td>
<td>($n = 13$)</td>
<td>($n = 13$)</td>
</tr>
</tbody>
</table>

$\sum_{x} = 82$

\[ \sum_{x_8} = 48 \quad \sum_{x_9} = 43 \quad \sum_{x_{10}} = 36 \]

The difference is again in the same direction as the difference for overlap error, but again it is not a significant one:

**TABLE 6(b)**

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form V, non-overlap form</td>
<td>1</td>
<td>17.551</td>
<td>17.551</td>
<td>3.389(1,72)</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>2</td>
<td>2.795</td>
<td>1.398</td>
<td>less than one</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Forms X Ages (Interaction term)</td>
<td>2</td>
<td>10.949</td>
<td>5.475</td>
<td>1.057(2,72)</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Within Cells</td>
<td>72</td>
<td>372.923</td>
<td>5.179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>404.218</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{1}\) For details of the calculation, see Appendix II, page iii.
All known structural errors are thus more frequent in the non-overlap version of the partially quantified three-term series problem. It is therefore clear that, unless executive errors are very much more frequent in the overlap form (and although I attempted to equate the problems for executive difficulty, it is necessary to check the differential effects of this error), the frequency of correct responses in the overlap form must significantly exceed the frequency of correct responses in the non-overlap form. Tabulation of the results confirms this:

TABLE 7

Correct responses (x) and executive error (y) in partially quantified three-term series problems

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>overlap</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\xi x)</td>
<td>49</td>
<td>58</td>
<td>84</td>
</tr>
<tr>
<td>(\xi y)</td>
<td>5</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>(n = 13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>non-overlap</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\xi x)</td>
<td>3</td>
<td>21</td>
<td>35</td>
</tr>
<tr>
<td>(\xi y)</td>
<td>5</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>(n = 13)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The frequencies of correct responses are shown as curves A and B in Figure 2. Over all, total correct responses are thus more than three times as /

1 Inspection of the responses confirms that there is no other answer which occurs consistently: it thus seems likely that all structural errors have been diagnosed. (Note that the converse would not necessarily have been true: if a response occurs frequently it might be due to a common executive error of the kind \(9 \times 6 = 72\).)
as frequent in the overlap form as in the non-overlap form - a result which is clearly highly significant without calculating an analysis of variance. It is clear also that the relatively small frequencies of executive error in these problem-forms could not significantly affect this result.

It will be recalled that, in order to avoid ambiguous formulation of three-term series problems (such as the use of the words 'difference' or 'between') it became necessary to ask two questions after the statement of the relationships: 'Who is older, A or B?' and 'How much older is he?' The first question was used not so much to gain information as to avoid the difficulties in formulation already described; the second question was the one intended to give information on overlap error and as such is the one which has provided the results already discussed.

The first question, however, does give an estimate of the children's ability to seriate the stated relationships:

\[
\begin{array}{ccc}
\text{TABLE 8(a)} \\
\text{Correct responses to the first question in partially quantified} \\
\text{three-term series problems} \\
\hline
\text{age 8} & \text{age 9} & \text{age 10} \\
\text{overlap form} & \frac{L_x = 51}{(n = 13)} & \frac{L_x = 64}{(n = 13)} & \frac{L_x = 77}{(n = 13)} \\
\text{non-overlap form} & \frac{L_x = 35}{(n = 13)} & \frac{L_x = 58}{(n = 13)} & \frac{L_x = 58}{(n = 13)} \\
\hline
\end{array}
\]

\[
\frac{\xi \xi x_{ov.} = 192}{\xi \xi x_{ov.} = 192} \\
\frac{\xi \xi x_{non-ov.} = 151}{\xi \xi x_{non-ov.} = 151}
\]

\[
\frac{\xi \xi x_8 = 86}{\xi \xi x_8 = 86} \\
\frac{\xi \xi x_9 = 122}{\xi \xi x_9 = 122} \\
\frac{\xi \xi x_{10} = 135}{\xi \xi x_{10} = 135}
\]
The results of an analysis of variance follow:

### TABLE 8(b)

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form V</td>
<td>1</td>
<td>21.551</td>
<td>21.551</td>
<td>3.660(1,72)</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>non-overlap form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>2</td>
<td>49.565</td>
<td>24.783</td>
<td>4.208(2,72)</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Forms X Ages</td>
<td>2</td>
<td>3.564</td>
<td>1.782</td>
<td>less than one</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>(Interaction term)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within cells</td>
<td>72</td>
<td>424.000</td>
<td>5.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>498.680</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thus while the results for this question are in the same direction as the results for the second question, the difference in this case between the two forms of problem is not significant. (There is a significant increase in correct responses with age; there is no significant interaction term.)

Although there was no significant difference in the frequency of overlap error between the fully quantified forms of the three-term series problem, it was decided to check these forms against each other with correct responses as criterion:

### TABLE 9(a)

1 For details of the calculation, see Appendix II, page iv.
Correct responses in fully quantified three-term series problems

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap</td>
<td>$\sum x = 39$</td>
<td>$\sum x = 89$</td>
<td>$\sum x = 70$</td>
</tr>
<tr>
<td>form</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
<tr>
<td>non-overlap</td>
<td>$\sum x = 51$</td>
<td>$\sum x = 63$</td>
<td>$\sum x = 75$</td>
</tr>
<tr>
<td>form</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
</tbody>
</table>

An analysis of variance yielded the following results:\^1:

**TABLE 9(b)**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form V non-overlap form</td>
<td>1</td>
<td>1.038</td>
<td>1.038</td>
<td>less than one (N.S.)</td>
<td></td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>2</td>
<td>88.693</td>
<td>44.347</td>
<td>5.237(2,72)</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Forms X Ages (Interaction term)</td>
<td>2</td>
<td>31.461</td>
<td>15.731</td>
<td>1.858(2,72)</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Within cells</td>
<td>72</td>
<td>609.692</td>
<td>8.468</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>730.885</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For details of the calculations see Appendix II, page v.
The F-ratio of less than one-eighth confirms the earlier finding that there is no significant difference in children's ability to solve the fully quantified forms of the problem.

Occasionally during the clinical studies there was a suggestion that children might add the two intervals in partially quantified problems when the relationship in the two clauses stated 'older than' and that they might subtract when the relationship stated 'younger than' (see, for example, the end of Rosalind's protocol - page 59).

If this is the case then we would expect a larger number of children to get the non-overlap 'older than' problems correct than the non-overlap 'younger than' problems; and we would expect a larger number to get the overlap 'younger than' problems correct than the overlap 'older than'. Similarly, we might expect a larger number of overlap errors in non-overlap 'younger than' than in non-overlap 'older than'; and a larger number of overlap errors in overlap 'older than' than in overlap 'younger than' problems.

The appropriate values are given in the accompanying table, and it can be seen that where the values are large enough to make comparison reasonably legitimate, only one of the entries bears out the hypothesis: age 9, non-overlap 'older than' compared with non-overlap 'younger than'. It thus seems that we can reasonably reject this hypothesis:

| TABLE 10 |

---

1 See also discussion and results on pages 149 ff. Note also that an analysis of variance was carried out on the results of those children who had fully quantified problems first. The results there were non-significant also.
TABLE 10

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>13</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-overlap/ 'older than'</td>
<td>2</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>non-overlap/ 'younger than'</td>
<td>1</td>
<td>8</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>25</th>
<th>30</th>
<th>41</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap/ 'older than'</td>
<td>25</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td>overlap/ 'younger than'</td>
<td>24</td>
<td>28</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>12</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-overlap/ 'older than'</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>non-overlap/ 'younger than'</td>
<td>8</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap/ 'older than'</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>overlap/ 'younger than'</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Criterion: correct responses (second question only)

Criterion: overlap error

It seems reasonable to assume that this hypothesis is related to the atmosphere effect hypothesis which is resoundingly confirmed in the case of two-term series problems (see next section). If this assumption is correct, it seems to be the case that at least in these instances, set effect works within premisses but not between premisses.

CONCLUSIONS /
CONCLUSIONS OF STUDY ON OVERLAP ERROR

It has been shown that children of ages 8, 9 and 10 produce significantly more overlap errors in their attempts to solve non-overlap (as compared with overlap) three-term series problems; and that this difference lies in the partially quantified rather than in the fully quantified forms. The same trend is shown by the two other structural errors which occur in these partially quantified three-term series problems, namely failure-to-respond, and the use of interval-values as answers. In these cases, however, the differences are not significant. As we would expect from the above results and from the low incidence of executive error, children achieve significantly more correct responses on the overlap form of the partially quantified three-term series problem than on the non-overlap form.

The set-effect hypothesis ('older than' = add, etc.) is not confirmed for between-premiss relationships in three-term series problems.

A comparison of the partially quantified overlap and non-overlap forms, problem by problem, using the three separate criteria of both questions correct, the second question correct, and overlap error, confirms in every case the findings reported above (see pages 146 ff.)
Asymmetry Error: Clinical Findings, Experimental Design and Results

In Donaldson's original investigation of three-term series problems it was noted that, in some cases, children made what she called 'asymmetry errors' (1963, p88). In such cases, children respond to a question of the sort 'A is 10; A is 3 years older than B. How old is B?' with the answer 13. In Donaldson's terminology, they treat "... an asymmetrical relationship as if it were a symmetrical one (ibid). Examples of this error were fairly frequent in the present investigation, and in many cases, the error seemed particularly persistent, confirming that it is structural.

It was further observed in the present investigation that asymmetry error seems to be much more frequent in fully quantified problems using the 'older than' relation than in those using the 'younger than' relation.\(^1\)

Now this observation can be explained ex post facto if we suppose that children tend to react to the phrase 'older than' by adding and to the phrase 'younger than' by subtracting. (Clinical evidence for this hypothesis is given below.) In these problems, as they stand, such a process would lead to asymmetry error in problems using the 'older than' relation, but not in problems using the 'younger than' relation; and this accords with what in fact did happen. The 'older than' form together with its asymmetry error is /

\(^1\) In fact, only one clear example of asymmetry error in a 'younger than' problem occurred and that was by no means persistent. Two other cases occurred but in these it seemed that the asymmetry error was a means of avoiding the overlap situation. (Thirty cases occurred in 'older than' type problems.)
is illustrated above; the 'younger than' form, which did not lead to asymmetry error is as follows: - 'A is 10; B is 3 years younger than A. How old is B?'

Graham's protocol illustrates both the structural nature of the error, and his painful efforts to preserve the rule 'older means add'. (Note that in his explanation, he was able to treat asymmetrical relations correctly, but not to subtract when the problem stated 'older than'.)

We want to find out the ages of Ian and David. We know that Peter is 8, and that he is three years older than Ian and one year older than David. How old are Ian and David?

S: 'Ian is eleven and David is twelve.

E: Good. Yes. How did you get that? Tell me how you worked it out.

S: I just added on three and that made Ian eleven, and then I added on one and that made David twelve.

E: How did you know to add on three?

S: It says Ian is ... We know that Peter is 8 and that Ian is three years older than Ian (sic) and I added on the three and that made Ian eleven. (This explanation of asymmetry error occurred fairly frequently. What the child seems to be saying is this: "It can't be Peter who is three years older than Ian, since Peter is eight and Ian is eleven - i.e. the child's answer - so the 'he' must refer to Ian." The child is thinking in terms of "Ian is three years older than Peter" but reads he is three years older than Ian, and so says "Ian is three years older than Ian." This may be an ex post facto attempt to justify his answer rather than the 'reasoning' which originally /
originally led to the answer, but it shows how far the child will go in order to preserve his asymmetry interpretation.)

E: Good, and how did you know that David was twelve?
S: I added on ... I added on another one.
E: How did you know to add on another one?
S: David is one year older than Ian, so I added on the one.
E: Than Ian, I see (i.e. a conventional overlap error combined with asymmetry.) Now listen to this. We know that Peter is eight and that he is three years older than Ian. Now if Peter is eight and he is three years older than Ian, how old is Ian going to be?
S: (Pause) Peter is eleven. (From now on the child seems to be fighting against the temptation to make Ian eleven, his original asymmetry error. He knows that Peter is older than Ian - see below - he must add; therefore Peter is eleven, since it can't be Ian.)
E: No, Peter is eight, isn't he? It says Peter is eight and that he is three years older than Ian, so how old is Ian going to be?
S: (Pause) Eight. (Graham still feels that he can't subtract since he's dealing with an 'older than' relationship; Ian is not eleven; Peter is not eleven; therefore Peter is eight and Ian is eight.)
E: Why will he be eight? (Pause) How old is Peter?
S: Eight.
E: Eight, and Peter is three years older than Ian, isn't he? So how old will Ian be?
S: (Pause) Eleven. (He seems now to have exhausted all other possibilities, and since he feels that he must not subtract, he returns to the classical asymmetry error.)
E: How do you know he'll be eleven? (Pause) If I said that Ian was three years younger than Peter, how old would Ian be?
S: (Pause) Five (i.e. 'younger than' has led to subtraction and the correct answer).
E: Five. Good. But it says that Peter is older than Ian so who is the older?
S: Peter.
E: Peter. Peter is older. Now he's three years older than Ian, so how old is Ian?
S: Five (i.e. when I built up from an unquantified situation, Graham was able to achieve the correct answer. It should be noted, however, that later in the same protocol Graham returned to asymmetry error once more).

On several occasions, in clinical interview, children have, simply and directly, identified the term 'more' with the process of addition and 'less' with subtraction.

In the following problem:

Jack weighs 7 stones. Jack weighs two stones less than Frank. What weight is Frank?

Lennox gave the answer '5 stones'. In reply to questioning, she said that she had subtracted.

E: How did you know to take away?
S: Because there's less (emphatically).
E: How did you know to take away when it says 'less'?
(No reply).
E: Does 'less' mean take away?
S: (Pause) Yes.

Although some element of suggestion on the experimenter's part may seem to be present towards the end of the interview it seems fairly clear that this is in fact the reason for the behaviour in question, so frequently do children point out the term 'less' as being the reason for subtracting (or 'more' for adding).

Summarising the historical description\(^1\), we have the hypothesis (based on clinical evidence, and seemingly supported ex post facto\(^2\) by the relative frequency of asymmetry error in different kinds of problems) that Donaldson's 'asymmetry error' is most often due to children reacting to the phrase 'older than' by adding and to the phrase 'younger than' by subtracting, irrespective of the formal requirements of the problem.

\(^1\) This discussion applies to fully quantified problems; asymmetry error in partially quantified problems, although not unknown, was very much less common.

\(^2\) Hence not usable as primary evidence.
Stated in another way, problem statements which include the phrase 'older than' seem to have an 'add' atmosphere and problem statements which include the phrase 'younger than' seem to have a 'subtract' atmosphere; just as, according to Woodworth and Sells (1935, p.453) an A-type syllogism "... has an all-yes atmosphere", and an E-type syllogism "... has an all-no atmosphere", etc. Rather than responding analytically to the formal requirements of the problem, these children seem occasionally to be "... influenced by the global impression of the problem" (Hunter, 1957(b), p.175). This phenomenon has been called the 'atmosphere effect', and it seems to be an adequate explanation, or at least description, of what is happening in this instance (see also discussion on pages 122 ff).

Now it is possible to design two kinds of problem, one of which (if our hypothesis is adequate) should lead to asymmetry error and wrong solution significantly more frequently than the other. Each type of problem has two forms: one 'older than' form and one 'younger than' form:

Type 1  
A is 10.  A is 3 years younger than B.  How old is B?  
A is 10.  A is 3 years older than B.  How old is B?  

Type 2  
A is 10.  B is 3 years younger than A.  How old is B?  
A is 10.  B is 3 years older than A.  How old is B?  

(In the first pair it will be seen that in the relational statement, the independent variable is stated first; and that in the second pair it is the dependent variable which is stated first.)

Now Type 2 is, according to our hypothesis, just as susceptible to 'atmosphere effect' as Type 1; but in Type 2, atmosphere effect should lead to the 'correct' solution. In Type 1, on the other hand, atmosphere effect should lead to asymmetry error and wrong solution. There should, in other words /
words, be a significant difference in the number of asymmetry errors occurring in the two types of problem-form detailed above, and this difference should be such that the number of asymmetry errors in Type 1 significantly exceeds the number of asymmetry errors in Type 2.

Since this form of the atmosphere effect does not formally require a three-term series for its occurrence - occasionally it occurred on one of the relations of a three-term series but not on the other - it was decided to test the hypothesis on two-term rather than on three-term series. In this way it was hoped to avoid unnecessary complications in the problem-form.

As stated above two kinds of two-term series problems were therefore constructed: one which, according to the hypothesis, should predispose to asymmetry error, and the other which should be free of asymmetry error. Within each kind, both 'older than' and 'younger than' problems were used:-

TABLE II

<table>
<thead>
<tr>
<th></th>
<th>older than</th>
<th>younger than</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>(Type 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-asymmetry</td>
<td>III</td>
<td>IV</td>
</tr>
<tr>
<td>(Type 2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here /

1 Where 'asymmetry errors' means 'wrong solution, presumably as a result of asymmetry error'. In every case group test results are used to infer the process which produced them, so strictly speaking, we can speak only of 'wrong solution' and not of 'asymmetry error' when referring to a numerical result. In order to avoid the rather clumsy statement 'wrong solution as a result of asymmetry error', however, it is proposed to substitute 'asymmetry error' for the phrase. Where it is necessary to distinguish between the process and the product - for example where the correct solution is achieved as the result of an error - the distinction will be stated explicitly.

2 As a result, of course, it is impossible to report on the interaction effects (if any) between three-term series difficulties and atmosphere effect but this sacrifice was felt to be justified.
Here is an example of each variety:

I Jean is 6. Jean is three years older than Sally.
What age is Sally? ______ years old.

II Jill is 5. Jill is two years younger than Susan.
What age is Susan? ______ years old.

III Jean is 6. Sally is three years older than Jean.
What age is Sally? ______ years old.

IV Jill is 5. Susan is two years younger than Jill.
What age is Susan? ______ years old.

From these examples it can be seen that asymmetry problems were exactly matched with non-asymmetry problems for content and for irrelevant structure. Twenty problems of each kind were constructed (ten 'younger than' interspaced with ten 'older than') and content and irrelevant structure were matched throughout. Each series of problems had interpolated identical 'spacer' problems, in an attempt to minimise 'set' effects between problems.

1 See complete sets of problems in Appendix III, pages xxviff.

2 Where 'asymmetry problems' means problems which according to our hypothesis are likely to predispose to asymmetry error, and 'non-asymmetry problems' means those which are unlikely to predispose to asymmetry error.

Note that the system of naming the asymmetry problems thus differs from the system of naming the overlap problem. Overlap problems were so named because they involved an overlapping structure (just as non-overlap problems involved a non-overlapping structure). Asymmetry problems, on the other hand, are so named because it was predicted (and confirmed) that they would predispose to asymmetry error (just as it was predicted, and confirmed, that non-asymmetry problems would not lead to asymmetry error. (Note, however, that the system of naming overlap error and asymmetry error is the same: both refer to a particular misinterpretation of a relationship, or relationships.
problems, and each series was permuted in cyclical fashion to give twenty versions of each kind. In this way two tests of forty items each were constructed - one composed of 'asymmetry' items and the other of 'non-asymmetry' items. (It was felt that, in order to avoid any possible interference effects no child should be given both asymmetry and non-asymmetry problems.)

The test was administered under the same conditions as those described for the 'overlap - non-overlap' investigation. The items were given to two matched groups of children - one kind to each group - at each of the ages 8, 9, 10 and 11. The children came from a single primary school and a random selection of two groups of 13 children was made from all children at each age level, equating the number of boys and girls in each group as far as possible. The cyclically rotated books were distributed to the selected children as they sat.

RESULTS

The numbers of asymmetry errors occurring at each age under each condition are given in the following table:

| TABLE 12(a) |

---

1 The spacer problems were not to be scored and hence were not considered when cyclical permutation took place.
### TABLE 12(a)

A comparison of the incidences of asymmetry error in two types of two-term series problem

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry form</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
<tr>
<td></td>
<td>£x = 44</td>
<td>£x = 62</td>
<td>x = 46</td>
<td>£x = 44</td>
</tr>
<tr>
<td>non-asymmetry form</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
<tr>
<td></td>
<td>£x = 12</td>
<td>£x = 8</td>
<td>x = 3</td>
<td>£x = 6</td>
</tr>
</tbody>
</table>

$\sum \Sigma x_8 = 56$, $\sum \Sigma x_9 = 70$, $\sum \Sigma x_{10} = 49$, $\sum \Sigma x_{11} = 50$

An analysis of variance on the figures from which Table 12(a) is derived yielded the following results:

### TABLE 12(b)

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetry versus non-asymmetry forms</td>
<td>1</td>
<td>268,163</td>
<td>268,163</td>
<td>41.441(1,96)</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>3</td>
<td>10,798</td>
<td>3,599</td>
<td>less than one</td>
<td>N.S.</td>
</tr>
<tr>
<td>Interaction (Forms X Ages)</td>
<td>3</td>
<td>10,029</td>
<td>3,343</td>
<td>less than one</td>
<td>N.S.</td>
</tr>
<tr>
<td>Within cells</td>
<td>96</td>
<td>621,232</td>
<td>6,471</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>910,221</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The /

1 For details of calculation see Appendix II, page vi.
FIG. 3. ASYMMETRY AND NON-ASYMMETRY GROUPS

Scale: 1 square = 3 units

CURVE E = NON-ASYMMETRY GROUP - CORRECT RESPONSES
CURVE F = ASYMMETRY GROUP - CORRECT RESPONSES
CURVE G = ASYMMETRY ERRORS - IN ASYMMETRY GROUP
CURVE H = ASYMMETRY ERRORS - NON-ASYMMETRY GROUP

AGES
The difference between the numbers of asymmetry errors occurring in each form of problem is thus very highly significant. (For the results to be significant at the 1% level, an F ratio of 6.96 is required.) There is no significant difference among the ages, nor any significant interaction term. (See curves G and H in Fig. 3 for a graphical representation of these frequencies.)

If asymmetry error were the result only of atmosphere effect, we would expect to find asymmetry errors only in asymmetry problems. Curve H, however, is a measure of asymmetry error in non-asymmetry problems, and since an atmosphere effect of the kind postulated above could, in non-asymmetry problems, only result in a numerical answer which coincides with the correct response, the cause of Curve H must have been other than the atmosphere effect. There is no clinical evidence so far concerning the nature of this very infrequent kind of response; a priori, however, it seems that this error is likely to be arbitrary in nature: there certainly seems to be no other psychological process possible in so simple a problem-form. (Executive errors, if randomly distributed throughout the likely range of values, are not sufficiently numerous to account for the incidence of these errors.)

Hence, like overlap error, asymmetry error seems to be a relatively infrequent error (the incidence of asymmetry error never exceeds 24%) but one /

---

1 An arbitrary response, one would imagine, would result from the child failing altogether to come to grips with the problems, and, seeing two numbers, arbitrarily adding or subtracting them.
one which extends over a wide age range\textsuperscript{1}. As in the case of overlap error, then, it seems that the structural difficulty which takes the form of asymmetry error is of a different kind from that which chiefly engages Piaget's attention (see pages 118 ff).

It has been verified that asymmetry error is likely to be due to atmosphere effect (as defined above). Now if atmosphere effect is the only source of difficulty in these problems (other than equated executive difficulty) then:

(a) the difference in correct response between types of problems should be in the direction indicated by asymmetry errors (that is to say, the problems prone to asymmetry error should have a lower incidence of correct response than problems not prone to asymmetry error);

and (b) the difference in incidence of correct response between types of problems should not differ significantly from the difference between types with asymmetry error as criterion\textsuperscript{2}.

If these differences are not as indicated, then some other factor must be operating, and further analyses will be necessary.

Using correct responses as criterion, the results are as follows:

\begin{table}[h]
\centering
\caption{Table 13(a)}
\end{table}

\textsuperscript{1} A developmental study is of course required to tell whether the same children continue to make the error from one occasion to the next - see Chapter 5. Meanwhile it should be noted that persistence of the error within an occasion in clinical studies (see, for example, Graham's protocol, quoted above) suggests that this might be the case. On the other hand, errors are spread fairly evenly throughout the asymmetry group and are relatively infrequent within subjects.

\textsuperscript{2} A discussion of the assumptions underlying this condition is given on pages 167.
TABLE 13(a)

A comparison of the incidences of correct responses in the two types of two-term series problem

<table>
<thead>
<tr>
<th>Age</th>
<th>Asymmetry Form</th>
<th>Non-Asymmetry Form</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \bar{x} = 125 )</td>
<td>( \bar{x} = 161 )</td>
</tr>
<tr>
<td>8</td>
<td>( (n = 13) )</td>
<td>( (n = 13) )</td>
</tr>
<tr>
<td>9</td>
<td>( \bar{x} = 170 )</td>
<td>( \bar{x} = 219 )</td>
</tr>
<tr>
<td></td>
<td>( (n = 13) )</td>
<td>( (n = 13) )</td>
</tr>
<tr>
<td>10</td>
<td>( \bar{x} = 295 )</td>
<td>( \bar{x} = 380 )</td>
</tr>
<tr>
<td>11</td>
<td>( \bar{x} = 295 )</td>
<td>( \bar{x} = 380 )</td>
</tr>
</tbody>
</table>

These results are represented graphically as Curves E and F in Figure 3. The difference between the two curves is shown by an analysis of variance to be statistically highly significant. The results of the analysis of variance are as follows:

TABLE 13(b)

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetry versus non-asymmetry form</td>
<td>1</td>
<td>347.12</td>
<td>347.12</td>
<td>14.96(1.96)</td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>3</td>
<td>639.73</td>
<td>213.24</td>
<td>9.19(3.96)</td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>Forms x Ages (Interaction term)</td>
<td>3</td>
<td>7.27</td>
<td>2.42</td>
<td>(N.S.)</td>
<td></td>
</tr>
<tr>
<td>Within cells</td>
<td>96</td>
<td>2226.77</td>
<td>23.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>3220.89</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 see Appendix II, page vii for details of the calculation.
Using correct responses as criterion, then, it can be said that asymmetry problems are significantly more difficult than non-asymmetry problems for children in the age range 8 - 11 inclusive. (It can also be said, although this is not the major issue, that there is a significant increase in correct responses with age for asymmetry and non-asymmetry problems taken together; and by inspection of the graph, probably for each problem individually, since their slopes are almost parallel. In the case of non-asymmetry problems, the slope levels off at the age of 10 where the graph approaches the maximum possible value; and in the case of asymmetry problems, levelling off begins between the ages of 10 and 11. It would be of interest to continue up the age scale with asymmetry problems to see when the maximum is reached and what the value is. The fact that the interaction term is not significant suggests that the relative difficulty of asymmetry and non-asymmetry problems does not change from one age to the next.)

Condition (a) is thus satisfied.

The differences at each of the ages 8, 9, 10 and 11 between Curves G and H are not however identical with the respective differences between Curves E and F:

<table>
<thead>
<tr>
<th>Table 14</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Difference between Curves E and F</strong>&lt;br&gt;age 8</td>
</tr>
<tr>
<td>$\geq x = 45$</td>
</tr>
<tr>
<td><strong>Difference between Curves G and H</strong>&lt;br&gt;age 8</td>
</tr>
<tr>
<td>$\leq x = 32$</td>
</tr>
</tbody>
</table>

In /
In an attempt to account for this difference an examination was made of the answers produced by the children. This examination indicated that no other numerical solution was consistently produced. It thus seems likely that the errors which are causing the difference are chiefly executive or arbitrary in type, probably the former.

The incidence of executive error in the two forms is given in the following table:

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry</td>
<td>$\leq x = 52$</td>
<td>$\leq x = 6$</td>
<td>$\leq x = 2$</td>
<td>$\leq x = 0$</td>
</tr>
<tr>
<td>form</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
</tr>
<tr>
<td>non-asymmetry</td>
<td>$\leq x = 51$</td>
<td>$\leq x = 1$</td>
<td>$\leq x = 1$</td>
<td>$\leq x = 1$</td>
</tr>
<tr>
<td>form</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
<td>$(n = 13)$</td>
</tr>
</tbody>
</table>

$\leq x_{\text{asy}} = 60$

$\leq x_{n-\text{asy}} = 54$

G.T. = 114

---

1 A count of failures-to-respond was also made, but since these are almost identical in the two treatments their effect is not a differential one and is not considered separately:
TABLE 16(a)

Incidence of executive error in the two treatments

<table>
<thead>
<tr>
<th>Age 8</th>
<th>Age 9</th>
<th>Age 10</th>
<th>Age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asymmetry Form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq x = 39$</td>
<td>$\leq x = 31$</td>
<td>$\leq x = 12$</td>
<td>$\leq x = 8$</td>
</tr>
<tr>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
<tr>
<td><strong>Non-Asymmetry Form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq x = 27$</td>
<td>$\leq x = 32$</td>
<td>$\leq x = 8$</td>
<td>$\leq x = 6$</td>
</tr>
<tr>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
<td>(n = 13)</td>
</tr>
</tbody>
</table>

Asymmetry problem-forms, then, have rather more executive errors and failures-to-respond than the non-asymmetry forms\(^1\) and as a result Curve B will be depressed, thus accounting for most of the difference.

Now it was claimed that when the problems were constructed, the two types were matched for executive difficulty. It thus becomes necessary to test to see whether the difference between the differences in Table 14 could have occurred by chance (as we expect) or whether they are likely to have been /

---

\(^1\) This might be partly accounted for in the following way (but see also (compatible) explanation on page 167): In the asymmetry-prone problems, there is a difference between the answer produced as a result of atmosphere effect and the analytically-produced correct response, whereas in the other problems, the answer produced through atmosphere effect coincides with the answer achieved by analytical means. In children who are in the process of changing from an 'atmosphere effect' response to the problem to an analytical approach, there is thus likely to be some confusion caused by the dissonant answers in the former case, whereas the answers reinforce each other in the latter. This dissonance might lead to confusion and hence to executive error.
been due to some real difference in the problem forms.  

An analysis of variance was carried out on the figures from which Table 16(a) was derived and the results are as follows:\footnote{For details of the calculation, see Appendix II, page viii.}

**Table 16(b)**

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetry form versus non-asymmetry form</td>
<td>1</td>
<td>2.779</td>
<td>2.779</td>
<td>less than one</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Between ages</td>
<td>3</td>
<td>87.644</td>
<td>29.215</td>
<td>2.812(3,96)</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Forms X Ages</td>
<td>3</td>
<td>3.567</td>
<td>1.189</td>
<td>less than one</td>
<td>(N.S.)</td>
</tr>
<tr>
<td>Within cells</td>
<td>96</td>
<td>999.539</td>
<td>10.391</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>1091.529</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

From the table it can be seen that, irrespective of our hypothesis, the difference in incidence of executive error in the two forms of problem (asymmetry and non-asymmetry) is not statistically significant. (Decrease in incidence of the error with age, although irrelevant to the present discussion, is significant.)
In spite of this, however, it is possible that executive error might reduce or eliminate the significant difference in frequency of correct responses which has been reported above. Curve F (asymmetry group, criterion correct responses) is depressed more by executive error than Curve E (non-asymmetry group, criterion correct responses) in Figure 3, and hence the gap between them is widened 'artificially' — that is to say, the difference, whose significance I have tested, is not due entirely to asymmetry error, but also to executive error.

If we take the differences between the frequencies of executive errors at each age, however, and add these values to the respective frequencies of asymmetry-group correct responses, we can eliminate the differential effect of executive error. The original figures, in which the asymmetry group scores are differentially depressed by executive error are as follows:

### TABLE 17

<table>
<thead>
<tr>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry form</td>
<td>( \leq x = 125 )</td>
<td>( \leq x = 161 )</td>
<td>( \leq x = 200 )</td>
</tr>
<tr>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
</tr>
<tr>
<td>non-asymmetry form</td>
<td>( \leq x = 170 )</td>
<td>( \leq x = 219 )</td>
<td>( \leq x = 248 )</td>
</tr>
<tr>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
<td>(( n = 13 ))</td>
</tr>
</tbody>
</table>

\( \chi^2_{\text{asym.}} = 694 \)

\( \chi^2_{\text{non-asym.}} = 884 \)

The
The modified figures, derived by the method of compensation described above are as follows:

**TABLE 16**

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>asymmetry form</strong></td>
<td>( x = 137 )</td>
<td>( x = 160 )</td>
<td>( x = 204 )</td>
<td>( x = 210 )</td>
<td>711</td>
</tr>
<tr>
<td></td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>asym.</td>
</tr>
<tr>
<td><strong>non-asymmetry form</strong></td>
<td>( x = 170 )</td>
<td>( x = 219 )</td>
<td>( x = 248 )</td>
<td>( x = 247 )</td>
<td>884</td>
</tr>
<tr>
<td></td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>n-asym.</td>
</tr>
</tbody>
</table>

When the two tables are compared, and when it is realised that the \( F \) ratio for the original figures was 14.96 as against a required 6.96 for the 1% level of confidence (and 3.96 for the 5% level) it is clear without a new analysis of variance, that the differential elimination of executive errors will not affect the results reported earlier. (If, instead of correct response, we take asymmetry error as criterion, we find that differential executive error could hardly affect these findings, except to increase the difference, and hence the significance: since executive errors are more frequent on asymmetry problems, they presumably act in the direction of reducing the number of asymmetry errors on asymmetry problems.)
One aspect of the results which has not yet received attention is the instability of the asymmetry error (and hence presumably of the atmosphere effect) within subjects. It seems that, over the whole age-range sampled, children are likely to respond sometimes as a result of atmosphere effect (and thus to make asymmetry error) and sometimes analytically. No child made asymmetry error consistently\(^1\). The largest number of instances in a single child is fifteen (the next largest is ten) out of a possible twenty, and even there the child was correct on three out of five of the other items. There appears to be no consistent relation with item-content, nor any with item-position (that is to say, if the child started, say, with item 9 - the items were rotated among the children - there was no tendency for asymmetry error to occur in this and in immediately succeeding items). On only one occasion did the combined frequencies of asymmetry errors, executive errors and failures-to-respond equal the maximum possible frequency, and on that occasion the child made one asymmetry error. On all other occasions at least two of the responses were correct. Thus, in spite of the several clinical examples of persistent asymmetry error within problems, the process within subjects between problems seems to be fairly unstable. The reasons for this instability (which extends throughout the age-range of our sample) will have to be sought by clinical study.\(^2\) Meanwhile it seems to provide further grounds for rejecting any attempt to match this process with any of Piaget's 'off-on' constructs.

\(^1\) On the other hand, the error was confined almost entirely to the asymmetry group and hence could hardly be an executive error. A longitudinal study aimed at the analysis of individual proneness to error has been carried out - see Chapter 5.

\(^2\) Subsequent clinical study has so far failed to indicate any reason.
(It should be noted that overlap error in three-term series problems behaved in rather the same way as asymmetry error, except that it appears to be confined to fewer children - less than half the group made overlap error. The asymmetry group, however, had twenty problems to solve, whereas in the case of overlap error, the group had only eight problems to solve. There was therefore less opportunity to make overlap error. On the other hand, more children made high overlap error scores, or - what may amount to the same thing - low overlap error scores and zero correct responses. In spite of this, however, a considerable proportion of children still made overlap errors combined with correct responses. It seems therefore, that overlap error is unstable in much the same way as asymmetry error, although to a lesser degree.

CONCLUSIONS OF STUDY ON ASYMMETRY ERROR

Statistical confirmation has been obtained that asymmetry error is due to a process, discovered clinically, in which children tend to add when the problem states 'older than' (etc.). I have identified this process with atmosphere effect. This hypothesis was tested by designing two-term series problems of two forms, one of which - if atmosphere effect (as defined) operates - would lead to wrong solution as a result of asymmetry error, and the other to correct solution. Using the difference in the frequencies of asymmetry error in the two problem-forms as criterion, the hypothesis was verified. (Curve H indicated the presence of a small proportion of asymmetry error which I assume was caused by an arbitrary manipulation of the variables.)

A/
A further comparison of the two problem-forms, using correct-response as criterion, indicated that the difference there was also significant and in the direction expected as a result of the differential incidence of asymmetry error. A direct comparison of the distributions - (G-H) with (E-F) - indicated that the differences in frequencies of the correct responses were rather greater than would be expected if only asymmetry error were responsible. A check of the children's answers for consistent responses among children indicated that it was unlikely that any other structural error was operating and that the difference (between the differences) was likely to be due to executive error.

An analysis of variance on the frequencies of executive errors indicated that there was no significant difference in their incidence; and adding the difference between executive errors totals to the appropriate values of the asymmetry group results indicated that the significant difference there was not negated.

Some tentative suggestions were made regarding the theoretical significance of asymmetry error, including the significance of its low incidence over the range of ages sampled, and its instability 'within children'. (See also chapters 5 and 7).
CHAPTER 5

Further Research on Overlap Error and Asymmetry Error; Validational and Longitudinal Designs and Results

On Pages 99-100 I pointed to the instability within subjects of overlap error and of asymmetry error. In the cross-sectional studies of the errors, the indications were (so far as we could see) that there was no differential liability to the errors; or in other words that no-one was more likely to make the error than anyone else.

Now it is possible to test for differential liability: if some children, in spite of inconsistency in making the error over a number of similar problems, are more prone to the error than others, then in a longitudinal study there should be a correlation between the two sets of scores which is greater than that expected by chance alone. In other words, if a child, in spite of his inconsistency, is more prone to the error than his neighbour, he will be likely to make the error again on the second occasion; if, on the other hand, there is no differential proneness, then there will be no such correlation between occasions.

Such a longitudinal study has been carried out, using as the first administration the cross-sectional study already described. At the same time, in order to check these results before treating them longitudinally the same problems were given to much larger (independent) groups.
groups at the same ages. We shall examine the results of the 'validation' study first:

Validation study

The overlap and asymmetry problem studies were originally carried out using relatively small groups (n=13) taken from schoolchildren in a Scottish Border town. (As we have seen the results, in spite of the small n, were highly significant.) The validation study repeated the experiments, using identical problem-forms and administrative procedure, but using much larger groups drawn from a number of schools in the City of Edinburgh. n ranged from 83 to 104 for the seven experimental groups.

The results of the validation study are as follows:

A. Asymmetry problems

Asymmetry errors /
### Asymmetry Errors

<table>
<thead>
<tr>
<th>Age</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>Combined Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>135</td>
<td>102</td>
<td>88</td>
<td>325</td>
</tr>
<tr>
<td></td>
<td>(n = 31)</td>
<td>(n = 29)</td>
<td>(n = 23)</td>
<td>(n = 83)</td>
</tr>
<tr>
<td></td>
<td>M = 4.4</td>
<td>M = 3.5</td>
<td>M = 3.5</td>
<td>M = 3.9</td>
</tr>
<tr>
<td></td>
<td>M = 3.6</td>
<td>M = 4.0</td>
<td>M = 4.8</td>
<td>M = 4.2</td>
</tr>
<tr>
<td>9</td>
<td>118</td>
<td>171</td>
<td>149</td>
<td>438</td>
</tr>
<tr>
<td></td>
<td>(n = 33)</td>
<td>(n = 40)</td>
<td>(n = 31)</td>
<td>(n = 104)</td>
</tr>
<tr>
<td></td>
<td>M = 3.6</td>
<td>M = 4.3</td>
<td>M = 4.8</td>
<td>M = 4.2</td>
</tr>
<tr>
<td>10</td>
<td>127</td>
<td>162</td>
<td>92</td>
<td>381</td>
</tr>
<tr>
<td></td>
<td>(n = 31)</td>
<td>(n = 41)</td>
<td>(n = 29)</td>
<td>(n = 101)</td>
</tr>
<tr>
<td></td>
<td>M = 4.1</td>
<td>M = 4.0</td>
<td>M = 3.2</td>
<td>M = 3.7</td>
</tr>
<tr>
<td></td>
<td>M = 4.1</td>
<td>M = 2.7</td>
<td>M = 3.1</td>
<td>M = 3.5</td>
</tr>
<tr>
<td>11</td>
<td>122</td>
<td>96</td>
<td>111</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>(n = 30)</td>
<td>(n = 35)</td>
<td>(n = 36)</td>
<td>(n = 101)</td>
</tr>
<tr>
<td></td>
<td>M = 4.1</td>
<td>M = 2.7</td>
<td>M = 3.1</td>
<td>M = 3.3</td>
</tr>
</tbody>
</table>

B. Non-overlap problems (partially quantified)

### Overlap Errors

<table>
<thead>
<tr>
<th>Age</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>Combined Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>17</td>
<td>26</td>
<td>26</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>(n = 32)</td>
<td>(n = 31)</td>
<td>(n = 32)</td>
<td>(n = 95)</td>
</tr>
<tr>
<td></td>
<td>M = 0.53</td>
<td>M = 0.34</td>
<td>M = 0.81</td>
<td>M = 0.73</td>
</tr>
<tr>
<td></td>
<td>M = 0.53</td>
<td>M = 0.81</td>
<td>M = 1.5</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>52</td>
<td>13</td>
<td>46</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>(n = 37)</td>
<td>(n = 29)</td>
<td>(n = 37)</td>
<td>(n = 103)</td>
</tr>
<tr>
<td></td>
<td>M = 1.41</td>
<td>M = 0.45</td>
<td>M = 1.2</td>
<td>M = 1.1</td>
</tr>
<tr>
<td></td>
<td>M = 1.41</td>
<td>M = 0.86</td>
<td>M = 1.5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>82</td>
<td>69</td>
<td>50</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>(n = 34)</td>
<td>(n = 29)</td>
<td>(n = 35)</td>
<td>(n = 98)</td>
</tr>
<tr>
<td></td>
<td>M = 2.4</td>
<td>M = 2.4</td>
<td>M = 1.43</td>
<td>M = 2.1</td>
</tr>
<tr>
<td></td>
<td>M = 2.4</td>
<td>M = 1.5</td>
<td>M = 1.5</td>
<td></td>
</tr>
</tbody>
</table>
By inspection, the majority of the means from the original study ($M_o$) are included within the range of the sub-means from the schools in the present study. In other words, the experimental classes might easily have fitted within the ranges of the control classes: e.g. at age 10, asymmetry error, the original mean is 3.5 and this is included within the range 3.2 to 4.1. The original mean can thus be accepted as being representative of the range of means in the larger sample.

The result outside the range in the asymmetry study is in a direction which would have made the results in the original study more rather than less significant: a more representative sample would have shown more, rather than less, asymmetry error. Our chief concern here is of course with patterns of error, and it should be noted that the discrepancy between the old mean and the lowest of the new means is only 0.1 of an error point. It appears that with some caution, therefore, we can generalise from the findings of the developmental study of asymmetry error for the eight year-olds; generalisation from the other age groups appears safe within the limits imposed by our larger sample.

The same cannot be said of the only non-overlap problem discrepancy. There the error discrepancy (at eight years) is in a direction which would have made the original results less significant\(^1\). Now if there were a large enough number of samples, we would expect an occasional 'rare' result as a matter of course. Since we do not have a sufficiently large number /

\(^1\)The original analysis of variance was recalculated on the assumption that the entry for age 8, non-overlap, should have been 9. (This estimate was based on the validation study assuming the original n=13.) Under these conditions, the Mean Squares were 16.615, 1.706 and 1.038. Assuming that the Mean Square within cells is not unduly altered - a reasonable assumption since five of the six cells upon which it is based are unaltered - then we can say that the conclusions of the analysis of variance are not different from those already reached.
number of samples to check this, I propose to generalise from this result, but with less confidence than from the others. (If it is thought necessary, of course, the results from age eight can be ignored without affecting the other longitudinal findings.) The fact that the longitudinal findings are the same for the eight year-old non-overlap group as they are for the other age groups suggests that we are safe in our inclusion of this group in the general results - although of course such ex post facto argument cannot be used as evidence without risk of circularity.

**Longitudinal study**

The two administrations, twelve months apart\(^1\) were given under similar conditions and each child received the same rotation of questions on the two occasions. The results were compared by product-moment correlation, the distributions of those approaching significance being subsequently transformed according to a formula suggested by Snedecor (1946, Pp. 45-46) for values under 10, and the coefficients recalculated.

One of the difficulties in interpretation is that the mean error scores on the second occasion for asymmetry error were considerably lower than on the first (and the mean correct responses much higher\(^2\)). One explanation might be practice effect, but in view of the interval between the/

---

\(^1\) This interval was considered as an appropriate one to avoid practice effects. (It should be recalled in this connection that there was no significant age trend in the relevant analyses of variance, and it seems reasonable to assume that we are safe in thus generalising from cross-sectional to longitudinal results.)

\(^2\) These have not been quoted since they are not directly relevant to the present study.
the administrations this appears to be highly unlikely. Whatever the reason\(^1\), however, for purposes of correlation it is the order of the scores which interests us, not their absolute values\(^2\). (As might be expected, a few of the children who had been tested on the first occasion were absent on the second — see Appendix II. There is no reason, however, to suppose that there is a relation between the absences and the variables considered here.)

For overlap error, none of the correlation coefficient was significant. In ascending order of age they were \(r=0.41(3)\) (d.f.9); \(-0.597(3)\) (d.f.9); \(0.31\) (d.f.10). This suggests that within this age-range at least, there is no consistency from occasion to occasion over the twelve month interval selected for the present study. For asymmetry error, none of the coefficients for (initial) ages 8, 9 and 10 was significant, but that for age 11 was significant at the one per cent level. (The respective coefficients were \(0.08\) (d.f.8); \(0.39(3)\) (d.f.10); \(0.33\) (d.f.8) and \(0.77(3)\) (d.f.9).)

It may be that this last result is an isolated one, due to the sort of chance fluctuation one expects in a number of samples. It should be noted however that an observed sample correlation of \(r=0.77\) would occur, with random samples of this size, less than once in a hundred times in the absence of real correlation between the variables in the population from which /

---

1. Individual development is also unlikely — see above footnote on age trends in the analyses of variance.

2. Note however that the error sources in asymmetry problems on the second occasion tend to be rather more uniform as a result of their low values than they otherwise might have been. The correlation coefficient may in consequence be reduced. This fact should be borne in mind when interpreting the results of this part of the study. On the other hand, in the case where the result was not attenuated on the second occasion (initial age 8) the coefficient was only 0.08, while in the overlap problems — where again there was no such attenuation — all coefficients were non-significant. It thus seems unlikely that attenuation will be a source of error in the case of the asymmetry problem results.

(3) Snedecor's transformation was applied in these cases.
which the sample was drawn; here it has occurred once in seven times. The alternative hypothesis would presumably be that there is a trend toward increasing stability of the asymmetry error, and hence perhaps of atmosphere effect, with age. (There might have been a similar trend with overlap error, had the age range included 11 year-olds.) The available evidence, in the opinion of the present writer, is not sufficient to resolve the issue.

Thus we can say that, at ages 8, 9 and 10 (initial ages) for both overlap and asymmetry errors, there appears to be no tendency toward differential proneness, that is to say, no tendency for some subjects to make the error more than others. For age 11 we leave this question open.
CHAPTER 6

Other Analytical Categories

Of the categories of error relevant to three-term series listed by Donaldson (see, for example, 'A Study of Children's Thinking', P113), two of the most frequent, asymmetry error and overlap error, have been discussed here in great detail. Of Donaldson's other categories, the second most frequent of all - arbitrary allocation - is relevant only to asymmetry problems, and there only marginally.

Perhaps because the problem-forms were simpler than those used by Donaldson, no case of attempting to calculate the value of the link-term occurred in fully quantified problems. In partially quantified problems, however, there were several attempts to give the link an absolute value (see pages 45 ff.) - apparently a related, if not an identical phenomenon.

The remaining two categories were relatively infrequent in Donaldson's work on three-term series problems, and at least part of their importance stems from the fact that they appear also (and more frequently) in problems of different kinds (ibid, P81). Of these, loss of hold appeared to diminish in incidence (as one would expect) as the problems in this study became simpler in structure.¹

Arbitrary /

¹ This is a clinical impression since there was no attempt at standardizing conditions in the 'thinking aloud' situation.
Arbitrary rule was most in evidence in the explanation: 'He is older because he is mentioned first': apart from this kind of instance, it made no appearance in these forms of the three-term series problem.

It should be noted that none of these other categories (except perhaps arbitrary allocation in asymmetry problems) is identifiable in group test from numerical results alone.
Conclusions and Discussion

A. The main emphasis in this investigation has been on the further study of two important structural errors - overlap error and asymmetry error - which children tend to make in their solutions to three-term series problems. These are the main conclusions which can be drawn from the present experimental work about each of these errors:

Overlap error

I. Recommendations concerning problem statement; clinical findings

1. Problem-statements should be as simple as possible:
   (a) numerically
   (b) in content (unless, of course, its effect is being studied as in décalages).
   (c) in statement: short simple sentences should be used; the link, in fully quantified forms, should be mentioned twice.

2. In the statement of partially quantified problems, the following points should be observed:
   (a) Successive statements should be in a single line. If the line is broken, it is important to avoid breaking it at a point which would encourage the child to treat a relative as an absolute: e.g. 'Bill is three years older than Harry'.
(b) Values should be stated in words, not numerals.

(c) The words 'difference' and 'between' should be avoided.

(d) In order to avoid the treatment of relatives as absolutes, the link should be mentioned first in each premise in overlap problems. (Although this did not reduce the statistical significance in group tests, there was a tendency to this error clinically.)

(e) The relative size of the first interval should be varied to avoid 'automatic' subtraction.

(f) Lines should be used in preference to dots to indicate the answer-location. If dots are used, their number must not vary among problems.

3. Overlap error was persistent in clinical investigation in overlap problems. Many older children made apparently structural errors while many younger children had apparently genuine success.

4. Clinically, overlap error in overlap problems seemed to resolve into 'inclusion' difficulties and/or 'return to the link' difficulties (see Appendix I, pages 157 ff.)

5. Partially quantified problems were often solved by 'converting relatives to absolutes'. There appeared to be two groups of children who did this: those who treated it explicitly as a conditional procedure ('If Jane were 10...'), and those who believed that they were able to derive absolute ages from the given. Children who adopted absolute ages for the link were not always able to see that a change in the value of the link would not affect the value of the difference they had found.
II. Statistical findings

Children in the age-range 8 - 10 inclusive were tested on two forms of the three-term series problem (overlap and non-overlap) and the following conclusions were reached as a result of analyses of their answers:

1. There were very few overlap errors in fully quantified problems of either form. This was probably because children are able to deal with such problems in two separate steps (see pages 41 f).

2. In fully quantified problems the incidence of correct responses did not differ significantly in the two forms.

3. In partially quantified problems overlap error was significantly more frequent in the non-overlap form. This was contrary to clinical expectation, and more study will have to be made of the reasons for this finding.

4. Overlap error did not change significantly with age and was low in relative incidence.

5. A 'validation' study, using larger numbers, confirmed the overlap error in non-overlap problems for ages 9 and 10, in preparation for the longitudinal treatment of these results. There was a discrepancy at age 8, and hence the longitudinal results for this age-group should be treated with some degree of caution. (A re-calculation of the original analysis of variance on an estimated value based upon the new findings for age 8 showed no significant difference in the original result.)

6. In the longitudinal study there was no significant correlation between the results of successive administrations for any age group sampled. It thus appears that at these ages there is no tendency for some subjects to be more prone to the error than others.

7. The relative incidence of other structural errors: failure to respond, and use of interval value as answer, was in the same direction as the earlier findings, but not significantly so.

8. Executive error was infrequent in partially quantified problems.
9. As we would expect from the above findings, correct response in partially quantified overlap problems exceeded correct response in partially quantified non-overlap problems, and this difference is obviously significant.

10. The relative incidences for the 'Who is older?' question are again in the same direction, but not significantly.

11. A possible atmosphere effect between the premises was not confirmed.

12. The results for the problems as a whole are confirmed when the results are examined problem-by-problem on three of the most important criteria.

Asymmetry error

I Clinical findings

1. The error is a persistent one, confirming that it is structural.

2. It appears to be caused by children responding to the phrase 'older than', etc. by adding and to the phrase 'younger than', etc. by subtracting. This appears to be an atmosphere effect.

II Statistical findings

1. Asymmetry errors were shown to be significantly more frequent in the asymmetry form of the two-term series problem than on the non-asymmetry version, for children in the age-range 8 - 11 inclusive. These two forms were evolved in an attempt to test the hypothesis that children react to the phrase 'older than' by adding (etc.) irrespective of the formal requirements of the problem, and hence this hypothesis was verified.

2. There was no significant change with age in the incidence of asymmetry error, and its incidence, although significant was never high.

3. /
3. Correct responses differed significantly in the two forms and in the direction expected as a result of the asymmetry errors. Most of the differences between the correct response results and the asymmetry error results was accounted for in terms of executive error; the very small residual was in terms of failure to respond.

4. Neither executive error nor failure-to-respond differed significantly in the two forms, nor did the addition of executive error to the appropriate asymmetry-form frequencies significantly affect the earlier findings.

5. The occurrence of asymmetry error was unstable within subjects (see also paragraph 7). An attempt was made to account for this in terms of existent ability masked by atmosphere effect.

6. As in the case of overlap error, a 'validation' study, using larger numbers, confirmed the asymmetry error results in the asymmetry problems for all ages sampled.

7. In the longitudinal study for initial ages 8, 9 and 10, there was no significant correlation between the results of the successive administrations. Age 11 was alone in producing a significant correlation and it was decided that this did not provide evidence sufficient for any conclusion to be drawn in this instance. Again it appears that (at ages 8, 9 and 10 at least) there is no tendency for some subjects to be more prone to the error than others.

More generally, it should be noticed that quite small (and logically non-significant) changes in the forms of problems appear to be psychologically highly significant when the problems are solved by children in the age-range of the present sample.

Finally, the results (and especially the overlap error results) show how important it is to check statistically findings derived clinically by the 'thinking aloud' method, before generalizing from overt to covert thinking. On the other hand, the clinical method has also shown itself to be of value as a source of fruitful hypotheses (and especially so in the case of the analysis of asymmetry error).
B. PROBLEMS IN INTERPRETATION OF RESULTS

(a) The significance of overlap error in overlap problems at the clinical level

One superficially puzzling aspect of the overlap/non-overlap study is the difference in incidence of overlap error in overlap problems in the clinical and statistical parts of the investigation. It will be recalled that at the clinical level, overlap error occurred frequently in overlap problems, whereas at the statistical level it occurred hardly at all.

What then is the status of overlap error in overlap problems? Are we to say that after all, and in spite of the considerable clinical evidence, we were wrong in classifying such errors as structural? Clearly there are no grounds for saying this. The negative evidence provided in the statistical test refers only to the conditions under which these results were obtained: in other words, to covert thinking. Rather does it seem to be the case that these structural errors were peculiar to the kind of thinking which has been called 'thinking aloud'. In other words, the processes involved in 'thinking aloud' were different in the case of the solution of overlap problems from the processes involved in covert thinking, and hence structural errors occurred in the former but not in the latter. (If the error had occurred in covert thinking but not in thinking aloud of course, the situation would have been very much more difficult. In tackling a situation of this kind we would of course be no worse off than the conventional investigator working without the 'thinking aloud' procedure.)

Now while it is necessary that there be a close parallel between the two kinds of thinking (thinking aloud and thinking covertly) if the method is to be fruitful in any particular case, it is no part of the hypothesis that there must /
must always be a close parallel. All that is required is that there is a close parallel sufficiently often to enable us to establish fruitful hypotheses concerning covert thinking rather more often than we otherwise could; the criterion is a pragmatic one. And in the case of asymmetry error, this certainly was the case; there a hypothesis was established clinically which was verified statistically at a very high level of confidence. The check on whether there is a sufficiently close parallel in any particular case is of course a statistical one.

(b) The possible theoretical relevance of differential structural error in overt and covert thinking

In the above discussion I maintained that in spite of its non-appearance in group-testing, overlap error in overlap problems, as it occurred in clinical studies, could still be structural. Why this structural error should occur in 'thinking aloud' but not in covert thinking is a question which cannot be answered in detail from the information which is available at the moment. Interestingly enough, however, this information provides an answer to the question raised by Neisser (1963) in the last paragraph of his paper:

"One further prediction is worth mentioning. It has become common to have subjects 'think aloud' during experiments on problem-solving... If the views expressed in this paper are valid, such a procedure will substantially change the nature and course of the thought process, by limiting it to the main sequence. It may be possible to test this prediction experimentally."

The present evidence provides the answer that in the case of children's solutions to overlap problems at least, the two kinds of thinking do produce different results.

Having /
Having established that the two can and do differ, however, it is necessary to go on and ask why they do differ in this case. An answer to this question, produced by further research, would perhaps throw some light on the critical differences between the two kinds of thinking and hence on the nature of covert thinking itself. To this end a comparison of the 'dissonance' of overlap error on overlap problems with the 'assonance' of asymmetry error on asymmetry problems in the two situations might help to elucidate the reasons.

C. FURTHER DIFFERENTIATION AND REINTERPRETATION OF ANALYTICAL CATEGORIES

(a) The refinement of the concept of structural error

The theoretical concept for which this experimental investigation has perhaps the greatest relevance is that of structural error. It has become clear, as a result of the study of both the overlap and asymmetry errors, that this concept requires a major extension.

Structural errors have already been defined as those which are the result of a failure to understand what is involved in the problem. Now in the case of each of these errors it was discovered in the course of the present study that their incidence did not change significantly over a fairly wide age range (see pages 31 f. et seq); it was further discovered that the incidence of correct response was fairly high at a young age and that the percentage increase in correct response over a period of years was slow.

It /

1 and presumably determined chiefly by a decrease in the incidence of executive errors.
It is not intended to argue that this kind of structural error is therefore unimportant. On the contrary, the incidence of asymmetry error in certain specified situations was very significantly different from the incidence of asymmetry error in other specified and superficially similar situations, and it can be surmised that this sort of error is often the cause of difficulty in arithmetical problems in normal school work. The same may be said of overlap error. But this sort of structural error can be contrasted with another kind: the kind where the child is completely unable to appreciate the structure of the problem for a number of years, and then quite suddenly, after a short period of fluctuation, becomes able to understand the structure from then on. (What I have elsewhere called the 'off-on' type of structural error.) The investigator whose work has been most directed to the discovery of this sort of structural error is, of course, Piaget, and the studies in which this type of error is perhaps most evident are those on conservation.

As an illustration let us take the example of a child who has not yet achieved conservation of amount, and who is given two equal balls of plasticine. Once he has agreed that there is the same amount of plasticine in the two balls, the experimenter takes one of the balls, rolls it into a sausage-shape and asks whether there is the same amount of plasticine in the ball as in the sausage. No matter how this question is phrased, the child has not yet reached the level of conservation of amount and will not agree that there is the same amount in the sausage as in the ball - he will even claim /

---

1 which marks the beginning of the stage known as 'Concrete Operations' - the stage during which various kinds of conservation, amount, weight and volume in that irreversible order, are achieved.
claim, that if the two were sweets, there would be 'more to eat' in the sausage-shape 'because it is longer' (or conversely, but less commonly, 'more to eat' in the ball-shape, 'because it is fatter' - but never seeing the compensatory relationship between the two statements, which is the basis of conservation). The 'structural error' then (although Piaget does not use the term) before the achievement of conservation of amount is to see amount changing with visual configuration. When the child has achieved conservation of amount he is able to understand that no matter what the changes in visual configuration, the amount of the material remains the same.

The important points for the present discussion are that Piaget claims that the change from non-conservation to conservation is (apart from a short period of fluctuation) a quite sudden and irreversible process, and that all children before the age of about seven years make these structural errors, since this is a necessary stage through which children must pass as their intelligence develops.

Perhaps one of the most convincing demonstrations of his claims about conservation is to be found in the application of his concept of 'décalages'. A child who can conserve amount, and has rejected the arguments concerning single-dimensions of the substance will use these same arguments in support of non-conservation of weight, until he achieves this too. A 7- or 8-year old, for example, will take two balls of plasticine and weigh them against each other, altering the amounts of each until they weigh the same. If the two balls are then removed and one of the balls rolled into a sausage, the /

1 although some of course reach the stage of conservation earlier than others.
the child will claim that, if they were replaced, 'the sausage would weigh more because it's longer' (or the converse as before). The naive argument about a change in visual configuration is thus not a result of the child's age alone but of the kind of content with which he is dealing. A structural error which would be rejected as childish with one kind of content is thus still accepted with another. And since each of these phases covers roughly two years (between achievement of conservation of amount and conservation of weight, and between conservation of weight and conservation of volume - where the same arguments are used again) we can say that to this extent at least this kind of structural error differs from the kind typified by asymmetry and overlap error, which appear to have a much longer life on the same content.

While verification of Piaget's findings awaits adequate longitudinal study¹ it nevertheless seems that the sorts of structural error which he has analysed do differ from those studied in the present investigation. Whether this difference is one of kind or degree remains to be seen: this distinction probably depends on the criterion used. For example, using as criterion our original definition of structural error, (see page 24) the two kinds of error are identical. Using as criterion universality/non-universality of the error prior to correct response (allowing Piaget's claims of 100% incidence for the moment) then the two are obviously different in/

---

¹ For further suggestions concerning the testing of Piaget's theory, see Pp. 134 ff.
in kind\(^1\). From the point of view of suddenness-of-change, it seems to be almost certain from the graphs of overlap and asymmetry error that there is no such sudden and irreversible change from error to freedom-from error. Again, if we can accept Piaget's claim in this respect then on this criterion, the two errors again differ in kind.

(b) Atmosphere effect

It is important to notice that the verification of the hypothesis that children tend to respond to the phrase 'older than' by adding\(^2\) has no direct bearing on the status of its interpretation in terms of atmosphere effect. It is not the applicability of the atmosphere effect hypothesis that has been verified, only that of the asymmetry hypothesis; conversely, if it is shown that the concept of atmosphere effect is inapplicable, this will not affect the relationship which has been established.

What /

\(^1\) This difference between Piaget's kind of structural error and the kind studied in this investigation may perhaps be explained in the following way (Piaget would almost certainly subscribe to this explanation of his type of error): - The structural error which has 100\% incidence is the result of the non-availability of function because the necessary (mental) structure has not yet developed; the low-incidence long-term error is perhaps the result not of the lack of development of any function, but of the habitual interference with this function by another. (This might account for the instability of asymmetry error, and to a lesser extent of overlap error, and for those children who in clinical study can be helped to free themselves of the error, and who subsequently return to it in spite of their earlier success.)

It should be noted, however, that no group tests of Piaget's experimental situations have yet been given (as distinct from statistical treatment of clinically administered tests). No exact comparison of Piaget's findings with ours is therefore possible. It may be that under these conditions, the differences between Piaget's structural errors and ours disappear.

\(^2\) referred to (as a temporary measure) in the following discussion as the 'asymmetry hypothesis'.

What we are doing is to attempt to tie the asymmetry phenomenon to a body of theoretical knowledge. The established relationship (the 'asymmetry hypothesis') is quite specific to the kind of problem used in this investigation. But it is also possible to construct more general hypotheses about this hypothesis, and 'atmosphere effect' is an example of this.

What is being suggested, in other words, is that the kinds of behaviour which define the theoretical concept 'atmosphere effect' are those which 'underlie' the processes which occur in the established asymmetry hypothesis.

Is the fit an accurate one? It will perhaps require further experimental investigation to establish this finally, but meanwhile an initial (and necessary) comparison of the two sets of information can be made. Sells and Koob (1937) give the following definition: "The 'atmosphere effect' is a temporary set of the individual, arising within a situation (e.g. problem), to complete a task with that one of several alternative responses (e.g. an inference or judgment) which is most similar to the general trend or tone of the whole situation (e.g. problem)" (P.516).

Woodworth and Schlosberg (1954) use "global impression" instead of "general trend or tone", as does Hunter (1958(b)) who adds "general feel" as an alternative description. Each of these definitions also implies or states that 'atmosphere effect' replaces analytic reasoning on the part of the subject, and in this latter requirement, there is obviously an accurate match between atmosphere effect and the processes which have been shown to underlie asymmetry error. But it is not so clear that children in the present investigation are reacting to the general trend or tone, global impression or general feel of the problem.

1 The fact that Chapman (1959) has since claimed that Sells' results can be accounted for in terms of "everyday" reasoning which is not sufficiently precise for syllogisms does not of course affect the general definition of atmosphere effect. (On the other hand it does suggest that we ought to advance the explanation 'atmosphere effect' tentatively since it is possible that, as in the case of Sells' results, a more satisfactory explanation for the error may subsequently be discovered.)
On the contrary, according to clinical evidence they appear to be reacting rather specifically, if non-analytically, to phrases like 'older than', while virtually ignoring the context in which they appear. On the other hand, they are responding to the atmosphere or 'feel' of these particular phrases, and moreover they are using the components of the whole problem in accordance with this feel. In the end, therefore, we come rather close to accepting the definition of atmosphere effect as applicable to the asymmetry hypothesis in this respect too. All that is required is an extension of the definition (which does not seem to alter the definition in any critical way) to include aspects of the situation as well as the general situation. The definition might then read, 'The 'atmosphere effect' is a temporary set of the individual, arising within a situation... to complete a task with that one of several alternative responses... which is most similar to the trend or tone (or non-analytic impression/feel) of the whole of or part of that situation'.

Arrival at such a definition does not of course mark the end of the task. In any particular instance of atmosphere effect it is necessary to say why subjects respond in this way. In some cases, the reasons for this appear obvious; in Woodworth and Sell's investigation into atmosphere effect in the solution of syllogisms, for example, they state the critical variables entirely in stimulus terms (an A-type syllogism "has an all-yes atmosphere" and so on). Hunter (1957(a)) in a reference to Burt's (1919) discussion, on the other hand, distinguishes two less obvious sources of atmosphere effect, 'direct statement' and 'inclusiveness'.

In the case of the processes underlying asymmetry error, it does not seem that any of these explanations is adequate. It is not immediately obvious /

1 But see previous footnote.
obvious why children should add to the stimulus 'older than'; 'inclusiveness' refers only to relationships between two premisses; and although 'direct statement' could be made to refer only to one premiss, there seems to be nothing of the kind that Hunter signifies when he uses the term. (There is no direct statement of the kind 'A is taller than B' from which the conclusion 'A is tallest' can be 'derived' - in the present instance the question concerns the absolute height of B, and there is no 'direct statement' to this effect.)

An answer to this question has been sought through further clinical analysis. Thus where a child says, "...because 'older' means 'add', the investigator has asked the question 'Why?' So far, attempts have been unsuccessful. It may be the case that in the course of their school work children encounter the phrase 'older than' chiefly in addition problems and 'younger than' in subtraction problems. An examination of school arithmetic texts and of teaching methods might confirm this hypothesis.

In the initial attempt to tie in this phenomenon with a body of theory the concept of set was used, and the name 'set effect' was suggested as a theoretical description of the process. Now many writers identify the atmosphere effect with set - Sells and Koob state: "'Atmosphere effect' is related to set ...": Miller, in Stevens (1951) states: "Atmosphere effect is another name for ... determining tendencies"; and Woodworth and Schlosberg state: "Presumably one becomes quickly set ..." - but the term set has several meanings, only one of which appears to coincide with atmosphere effect. (English and English (1958) give three quite distinct definitions.) It /
It was thought better, therefore, to adopt the more restricted term 'atmosphere effect' especially since the generality of the term 'set' did not seem to add anything to the meaning of 'atmosphere effect'.

(c) Note on atmosphere effect and Piaget's structural errors

The relationship between atmosphere effect/asymmetry error and Piaget's discussion of children's ability to handle transitive asymmetrical relations (see for example Piaget, 1953, PL; 1952, chapter 6) is an interesting one especially in view of the earlier discussion on structural error. For Piaget, these relations form one of the eight elementary groupements and hence their development is a basic condition for concrete operational thought - that is to say, according to Piaget, the average child of seven should be able to handle asymmetry relations adequately and consistently. We find, on the other hand, that children make errors in their handling of such relationships until they are at least eleven. What is the explanation of this discrepancy?

The explanation seems to lie in a distinction which can be drawn between a child's actual ability to handle a relation, and whether, under certain conditions, he actually makes use of that ability or not. The conditions under which Piaget tests for the ability to handle asymmetrical relations are such that, if the ability is present, it is highly likely to be observed; to this end he has modified his test situations from predominantly verbal ones in the early days of his researches to situations in which the child has to handle actual materials in his more recent work. We have studied the ability to handle asymmetrical relations in problem-forms which are purely verbal, however, and under these conditions it has become apparent that children /
children sometimes fail to use their ability to handle asymmetrical relations and instead respond to some rather superficial aspect of the problem.

At least two facts are concordant with this interpretation. First, in group tests asymmetry error is, as we have seen, rather unstable and second, in the clinical situation, a child who has made asymmetry error in a problem may with help rid himself of it and achieve the correct response spontaneously and analytically.

If this interpretation is correct, these asymmetry errors are a product not so much of the relation with which the child is dealing but of the non-concrete form of presentation of that relation - an interpretation which is consonant with my identification of the process underlying asymmetry error as atmosphere effect.

D. THE NEED FOR STATISTICAL VERIFICATION

One of the chief lessons of the overlap part of the study is on the need for statistical verification of findings derived clinically. A good deal of the work carried out at the clinical level in the study of overlap error might have been more effective had an earlier comparative test of overlap as against non-overlap problems been made. The lesson seems to be this: not only should "...clinical studies come first and guide the statistical ones" (Donaldson, 1963, P.34) but there should be an interweaving of the two processes throughout the length of the research. The earlier formulation of testable hypotheses has its dangers^1 of course: it will require great flexibility /

---

^1 but also its advantages: an early group test carried out as soon as the clinically-derived ideas are reasonably well formulated could help to estimate the relative frequency (and hence to some extent, the importance) of the processes to be studied.
flexibility on the part of the experimenter to enable him to change these hypotheses when necessary. And the clinical phase will still be the earliest of all - as Donaldson points out, our own preconceptions often prevent us from making fruitful hypotheses a priori (ibid, P.33). But the proposed yet closer liaison between clinical and statistical techniques should help ensure that the hypotheses achieved in this way are more certainly fruitful than those achieved by clinical techniques alone.

B. BEYOND CLINICAL/STATISTICAL METHODOLOGY

The chief value of Donaldson's (1963) study seems to the present author to lie in its methodological suggestions. She gives valuable guidance on the conduct of research - guidance which as we have seen is in accordance with contemporary ideas - but says little or nothing about topics for research. The value of her method depends upon the use to which it is put.\(^1\)

It/

---

\(^1\) Donaldson actually used the method to reorient intelligence test construction but it is important to realize that this is not its only use. I feel that one use of particular value seems to be to revitalise educational research of the classroom-learning kind. There is no current failure to realize the need for such educational research - see for example the new 'Journal of Research in Science Teaching', first issued in 1963 - but research is usually of the arid traditional kind, e.g. Scott, N.C. (1964) in this journal. What we require is not so much the direct application of theory in terms of predictions from that theory to the classroom learning situation, as the use of theory in orienting clinical observation and experimental verification (Piaget's theory of cognitive development for example should not be used for abstract predictions about say science teaching to the classroom situation, but should instead be used to guide clinical observation of children's actual classroom behaviour in science lessons). That the Russians are making progress in this direction is indicated for example by the article by Menchinskaya: 'Some aspects of the psychology of teaching' (1954) in which clinical research into concept formation in geometry and geography is briefly described. Simon (1963) editing a book of readings from the U.S.S.R. in educational psychology writes "Psychologists therefore are closely in touch with educational developments, indeed directly concerned in planning them" (P.5).
It seems to the present writer that something more than this is required. Psychology has often, and especially recently, appeared to proceed in diverse and unrelated areas of the field, guided more by its methods than by any coherent aims. One author, Taylor (1958) went so far as to suggest that experimental method had become a cloak for intellectual sterility.

That a certain uneasiness about aimlessness is felt in the U.S.A. is perhaps best illustrated by the publication (from 1959 to date) of the American Psychology Association's Project 'A', a seven-volume work entitled "Psychology: A Study of a Science", whose aim is to review the present position with regard to content as well as method. (The U.S.S.R., as we might expect, does not have these problems to the same extent. Smirnov (1955) director of the Institute of Psychology of the Academy of Educational Sciences, reviews in Voprosy Psikhologii psychological research for the years 1953-1955. In this article, in his role as director, he is often severely critical both of particular papers and of the general direction research is taking, e.g. "...In developmental psychology, a serious deficiency is the small amount of work devoted to the upper age ranges; such work as there is relies more on isolated observations than on broad factual material, systematically collected and leading to further investigation ... ").

Obviously there is more than one possible solution to this problem of 'aimlessness'. Project 'A' represents one such attempt. Another might lie in something which has been advocated (usually in vain, as far as psychology is concerned) by philosophers of science for a number of years: the raising of a coherent body of theory on a basis of experimental findings, but /
but a body of theory which goes beyond these findings and from which predictions, not directly related to the initial findings, can be made; the testing of these predictions, and the consequent verification of the theory or its modification, or its abandoning.

What I suggest, both as one possible solution to current problems of direction, and even more as it appears to me to be intrinsically a worthwhile undertaking, is that we take Piaget's theory of cognitive development qua theory, and test it as such. Several suggestions for such an undertaking are given below. Meanwhile it might be wise to try to anticipate possible objections to such a programme.

Some possible objections

What I am proposing in the way of theory testing has little to do with attempts at constructing general psychological theories which were characteristic of the 1930's and 40's. It is this sort of general theory which Koch (1959) has in mind when he suggests that the Age of Theory is passing. While there is no doubt that, as Drever (1961) says, we must "...go back to our data with that puzzled humility which is the only frame of mind appropriate in the biological field..." this in no way implies that we must give up theory-making altogether. Piaget's theory is after all restricted to the field of cognitive development and thus meets these objections.

Hebb /

1 A plea for the construction of a theory of this kind, without direct reference to Piaget, has been made by Bruner (1957) - see below - and by Festinger (1957).
Hebb (1958) makes the point that, at the present stage of theory building, it is healthy that one theory should replace another without full verification. I would like to agree, but make a provisional exception of Piaget's theory if only because of its careful basing in clinical observation and its relative comprehensiveness within the field of cognition.

On the question of what kind of theory it is worth considering, Atkinson (1963) suggests (P.146) that qualitative models are inadequate and that quantitative (mathematical) models are necessary from the start. In case Hebb's criticism of mathematical models (see above, P. 12) is not considered sufficient to invalidate Atkinson's objection, it might be well to look at the reasons for it. He claims first of all that "... a theory based only on qualitative distinctions leads to a small number of testable predictions". If, like Atkinson's, a theory deals with a very limited range of behaviour - his deals with discrimination learning - we must agree. If the range is broader, then there appears to be no a priori reason for making a statement like this, although subsequent quantification will almost certainly be required to encompass individual differences. A case can however be made for approaching human behaviour qualitatively first. (Bruner (1962) would go even further (too far in the opinion of the present writer) when he says that one can do worse than live with a metaphorical knowledge of such processes for a number of years.)

Atkinson's other point is that "... the absence of precise systematisation usually leads to pseudo-derivations from the theory - that is derivations which require additional assumptions that are not part of the original theory... An important function of the mathematical model is to clarify/
clarify this aspect of theory." Although this danger is not an inevitable one, it is certainly one which anyone who attempts formal tests of Piaget's theory should guard against. (But see F.141 ff. for an important additional safeguard inherent in one of the proposed methods.)

A methodological objection in the area of theory testing comes from Hebb (1959) who says that "...Logically one cannot test the validity of a theory as a whole if it is based on a number of separate postulates and (in effect) consists of a network of more specific hypotheses. The experimentum crucis idea applies only to the single hypothesis of the type that links two independently variable entities." (Sparkes (1962) makes a related point, taking Popper to task on whether theories are infirmed by negative instances. He suggests that Popper's claims are based on well-established theories (e.g. gravitation) and says that theories are rather modified and added to, and ideally included in a more general theory, a point I take account of in the present paper). Again Hebb (ibid, P.627) says "Surely in our theorising we work by a series of approximations very rough in the early stages. As we find evidence that we are getting closer to the bull's eye, as theory develops, much more precision of statement and formulation will be in order." In answer to Hebb's first point, I would suggest that since Piaget's theory is more than 'a network of more specific hypotheses' we can test more than relations between independent variables, and would cite Piaget's operational structures (e.g. concrete operations) as an example (see below). (At the same time it must be admitted that we cannot test the whole theory in one critical experiment, but I do not make this claim.) With regard to Hebb's second point, I would plead as before that the existence of a complex theory of this kind requires that we do something to test it.
Other advantages of the proposed programme

Another feature of Piaget's theory which makes it worth testing is that it appears to give broader theoretical orientation to work by several other psychologists within the area of cognition: e.g. Miller, Galanter and Pribram's TOTE unit (1959), the concept of negative feedback generally, Festinger's (1957b) cognitive dissonance appear all to be related to Piaget's mobile equilibrium ¹ and in fact may be merely special cases of it since Piaget's concept is more general, including as it does both assimilation - distortion of the stimulus to fit with preconceptions - and accommodation - change of existent cognitive structures to 'account for' new experience. (The others appear to incorporate the notion of accommodation only.)²

It also seems highly likely that Goldstein and Scheerer's (1941) concrete and abstract thinking dichotomy can be related to Piaget's developmental, more highly differentiated, and hence more useful distinction between concrete and formal operations (although here - as in others of the examples given - there is no doubt that Piaget's system might also benefit from such an association).

In "The Growth of Logical Thinking" (Inhelder and Piaget, 1958), Piaget is of course defining the psychological aspects of Hebb's autonomous central processes which Hebb defined chiefly at the physiological level. The work of Hebb and Piaget can be seen as complementary, although in the growth of early learning, Piaget's emphasis on the role of behaviour provides a corrective for Hebb's possible overemphasis on perceptual learning. (It is /

¹ It would appear to be worth tracing its relation too to the social equilibrium described by Homans in his 'Human Groups' (1950).)

² In the motivating incongruity which Piaget describes between assimilation and accommodation, there may be a relationship with the sort of incongruity described in Harvey (1963) - ed. - "Motivation and Social Interaction", especially in the chapter by Hunt.
is interesting to notice that Bruner (1957) attempts a similar synthesis of divergent fields of study, by using a single conceptual framework—cognitive activity as a 'coding system'—and interpreting the work of others in these terms. One advantage of a synthesis of this sort would be the resolution of difficulties in terminology, and such difficulties occur even within American cognitive psychology: see e.g. Osgood's (1957) account of his difficulties in understanding Brunswik at the Colorado Symposium (P.33) in 'Contemporary Approaches to Cognition') "And there is a translation problem: we all use rather different languages for talking about cognition. In many cases I am sure that adequate translations are going to reveal that we are saying approximately the same thing in different terms..." (The drawback of course is that Piaget's language, although the most comprehensive is also the most difficult to acquire!)

Suggestions for a testing programme

There have of course been many attempts to replicate Piaget's findings (see e.g. Flavell, 1963, Chapter 11; also Lunzer's 1962 review), but such studies can hardly be considered as critical tests of the theory. It can in fact be argued that not only are they not sufficient but that they are not even necessary. For if predictions from the theory are verified then such verification in itself is sufficient, and verification of the initial findings upon which the theory is based become unnecessary (unless of course we wish to establish norms for such behaviour per se).
Even Piaget's own 'validation' work does not seem to be conceived in the way that I have suggested. When he speaks about testing his theory, he talks of "...controls both statistical and non-verbal" (1963). The same may be said of his principal co-worker, Inhelder (1962, Pp.21-22).

I am arguing for a full and explicit treatment of Piaget's theory qua theory, its testing, and subsequent verification, modification or abandoning.

My first set of suggestions for testing Piaget's theory use the traditional technique (outlined above) of making fresh predictions from the theory and of testing these predictions against reality. And second, I want to suggest a new and what I consider to be a potentially more powerful way of testing the theory.

1. (i) Piaget's theoretical concept 'internalisation' appears to be open to experimental test (as indeed every postulate in an adequate theory must be). According to the theory, thinking develops as a result of the gradual (developmental) internalisation of initially overt behaviour. Thus until 18 months or 2 years, according to Piaget, all activity in the child is overt. After that age, a gradual process of internalisation begins which by the age of 4 or 5 results in the ability to represent actions and events which could previously only be overtly performed. (Progressively more efficient manipulation of these internalisations, according to Piaget, results in the formation of Concrete Operations around the age of 6 or 7 years.)

Now /

---

1 Psychologists in the U.S.S.R. have, apparently independently, come up with what appears to be an identical explanation of the genesis of thought processes (see below).
Now according to Smith et al. (1947) the progressive and finally complete paralysis of an adult subject by curare has no observable effect on his central processes as measured by EEG throughout the experiment, and according to communication between the subject and experimenter prior to complete paralysis. (A series of pre-arranged muscle-twitches was used for communication as complete paralysis was approached.) Retrospective reports tended to confirm normality of central function (as far as it could be assessed in this way). Now if we can accept this experimental evidence, then it suggests that internalisation (if it occurs) is complete in (at least some) adults. Evidence for the concept of internalisation would therefore be provided if children's thinking deteriorated when muscular movement was prevented in a similar way. For obvious reasons, this programme is hardly feasible. An alternative but related programme however appears possible: Osgood (1953, Pp.648 ff) reviews several studies designed to test the role of motor activity during thought in human subjects. None of the experiments reported used a developmental approach (nor does Osgood suggest one) but it seems that such a developmental extension would in fact provide the test we are looking for.

In these experiments it was shown that muscle movements appear to occur in adults during thinking which concerns bodily movement. (That these are in fact not simply "...a general overflow of motor tension during mental work" - ibid, P.650 - is shown by the fact that in thinking of raising the right arm, action currents are generated in that arm but not in the left - Jacobson, 1932). Now if these muscle movements are semi-internalised actions (or, as seems more likely from Smith's experiment reviewed above, if/
if they are merely residuals, the result of internalisation) it should be possible to demonstrate a negative correlation between the amplitude of such movements and age. Such a finding would provide verification of Piaget's concept of internalisation; failure would throw serious doubt on the concept. That there is a negative correlation between amplitude and I.Q. in adult subjects (Osgood, Ibid, P.651) suggests that we may in fact get evidence of the sort which would support Piaget's theory from such a study.

(ii) Another test of the concept of internalisation is suggested by work of Slavina (1954) in the U.S.S.R., work not formally connected with that of Piaget. In reporting remedial teaching experiments with children who had failed to achieve a concept of number, Slavina claimed that it proved impossible to assist unsuccessful pupils without a gradual process of internalisation. If they went from the manipulation of objects straight to tasks without objects there was "...a deleterious effect on the quality of mental processes" (ibid, P.207). But where internalisation was assisted (e.g. the hand movements initially involved in counting were gradually attenuated until none remained; and counting aloud was similarly attenuated through whispering until silent counting was achieved), then children proved successful. A formal test with matched groups, only one of which has specific help with internalisation, should provide a critical test of this hypothesis.

2. Another possible test of Piaget's theory is to be found in so-called 'learning' experiments. Such experiments attempt to discover whether experience /
experience of a problem-specific sort is sufficient to enable the child to progress to a new level in his thinking (e.g. from concrete to formal operations) in that particular problem area. A large number of studies of this kind have been carried out (see Flavell, 1963, Chapter 11), but none of the authors appears to see the possibilities in such a design of a test of Piaget's theory. Their aim, according to Flavell (ibid), who has reviewed experiments of this kind to 1962, has been merely "...to find out what sorts of experiences do and do not facilitate development of the concept under study."

As I have implied, these designs can however be used not only to provide information concerning the acquisition of cognitive structures, but also as a test of the theory. The test which I suggest is twofold:

(a) Can structures be developed by deliberate teaching? (For this part of the experiment we can use any or all of the experiments described in Flavell (ibid).) If the answer is negative, then the theory stands. If positive, then we must ask a second question:

(b) Is this learning generalisable? If so then Piaget's theory again stands. If not this provides a critical disproof of this (a major) part of the theory, since, if the behaviour had not generalised and yet were firmly established, and within its own content area observed the appropriate rules for its particular kind of structure d'ensemble (e.g. groupement) then this would be evidence against Piaget's theory. (According to the theory - see e.g. Piaget, 1957 - the rules for operations in a particular content-area are not specific to that area but - when fully established, e.g. when operations /
operations are reversible - are general and independent of content. Décalages are of course a fly in the theoretical ointment: if the behaviour in question does not generalise, then Piaget might say that this was a case of décalages - c.f. conservation of amount but not of weight at ages 6 or 7. To avoid this theoretical Morton's Fork, the operation of décalages requires accurate definition so that an independent rather than ad hoc decision can be made in any particular case.)

As it happens, the available results from past experiments (although they had a different aim) go much of the way to providing the evidence we need since the results in almost every case are negative. "Probably the most certain conclusion is that it can be a surprisingly difficult undertaking to manufacture Piagetian concepts in the laboratory...most of them have had remarkably little success in producing cognitive changes" (Flavell, 1963, P.377). Moreover, Flavell adds that "...there is more than a suspicion from the present evidence that when one does succeed in inducing some behavioural change through this or that procedure, it may not cut very deep" (P.377). In support of this rather sweeping statement he quotes the extinction experiment by Smedslund.

With reference to the proposed design, then, we can say that since the results at stage (a) are so far negative we cannot at any rate at the moment go on to ask the second question. Up to this point the theory appears to stand the test.

3. /

---

1 See pages 129ff for further description of this example of décalages.

2 In this experiment Smedslund comes near to making a test of the sort I am suggesting.
3. Critical tests are also provided by Piaget's concept of 'synchronisms' (see "The Growth of Logical Thinking", 1958, Pp.264-265). Piaget claims that evidence for the existence of general structures, not necessarily consciously accessible to the subject, is to be found in synchronisms - the fact that formally related behaviours occur at the same age although the subject himself is not aware of their relation. As Piaget points out, for example, 9 - 10 year olds become able to tackle his balance-scale problem and toy dumping truck problem at the same time because these tasks depend for their solution on the same underlying structures. "In either situation the child understands that the force exerted by the weight changes as a function of the spatial relationships... A great many examples of the same type could be furnished from the concrete level. At the formal level, the synchronisation of like reactions in the face of analogous problems is still more striking" (ibid, P.265).

So far as I am aware, Piaget has remained content with noting such synchronisms as they occur. At least two critical tests of this concept seem possible. One is to predict certain behaviour as a result of a knowledge of the supposed underlying structure (allowing as before for decalages), and then to test our prediction against experimental observation. (If they do occur, this strengthens the theory; if not the theory must of course be modified). A second test is more obvious. It is based on the assumption that underlying structures should facilitate the same content abilities in different subjects, assuming also similar past experience in these content-areas. The test then consists simply of a correlation of abilities: If the emergence of ability A is correlated with the emergence of/
of ability B in one subject, and this is explained according to the theory as a synchronism, then is this correlation observed in all subjects tested?

4. Another and quite different way of testing Piaget's theory of cognitive development is to be found in computer simulation of human behaviour\(^1\). The majority of such studies (e.g. Newell, Shaw and Simon, 1957) have remained at the level of attempted simulation of particular behaviour. Newell et al., for example, attempted to simulate the behaviour of Whitehead and Russell in writing a section of 'Principia Mathematica'. Recently, however, there has been a move towards the use of computer simulation as a test of theory. Uhr (1963), in suggesting this approach, also claims certain advantages for this particular method of theory testing:

"...computers present the unique opportunity to explore efficiently and effectively with any theory or model, subject only to the restriction that the theory be adequately stated. And in this restriction itself lies a major virtue of the computer for psychological research. For the process of getting a model on to the computer is the process of describing the model precisely...The successful running of a programme put into the computer's language is, at the same time, a statement of the theory and a test of its consequences." (P. 232)

Since Piaget's theory is a developmental one, it would obviously be best to commence by simulating the behaviour of a newborn infant, starting with reflexes and moving on to the primary circular reactions (the first 'associations' /

---

1 Since writing this, it has come to my attention that a similar suggestion has been made by Simon, working on the computer simulation side of the fence. This suggestion is made in a note following an article (written for another purpose) on computer simulation and included at the end of the conference on the Intellectual Development of Children (1962). The fact that such a suggestion should be made by one specialising in computer simulation reinforces considerably the suggestions made here.
'associations' learned by the child) etc., described in Piaget's 'Origins of Intelligence in Children' (1953(b)]. Analogies to the reflex forms of behaviour, and to the stage-independent processes of assimilation and accommodation, would have to be built into the computer. We would then give the computer the kinds of experience (metaphorically speaking) out of which, according to Piaget, the primary circular reactions are built, and see whether Piaget's theoretical statements give sufficient information for the development of these and subsequent cognitive structures. If they do, then this will be a vindication of the adequacy of the theoretical system; if not then ways in which the theory must be modified should be apparent.

On the feasibility of such a programme, we have the statement by Reitman (1963) who, talking of the possibilities in computer-simulation generally, says (Pp.91-92): "Given the flexibility of the computer, however, it is extremely difficult to conceive of a structure or process which could not be represented, at some cost, or to some degree of approximation..."

It might be argued that there is no point in using anything as complex as a computer for simulating simple conditioning at the early developmental levels (and not say a simple learning-machine like Gray Walter's 'Machine Docilis' (1953).) My argument for doing so is two-fold. There is first of all the simple one that we are hoping eventually to simulate complex adult cognitive processes and the experience gained at the simpler level will obviously be invaluable at the higher levels of simulation. Not only this, /

---

1 In doing so, we would be at an advantage over conventional computer simulation studies which usually attempt to simulate complex adult behaviour, a more difficult task - see also below.
this, but Uhr (1963) has found that in the recognition of patterns (e.g. cartoons, printed and written letters, spoken speech) the computer did better if it were allowed "... to develop its own set of operations through 'experience'" rather than having these built in." Secondly the developing human being, whose processes we are simulating starts with a complex brain even for his initial simple acquisitions. And this is not merely an academic point. Hebb (1958, P.453) makes the same point in a different context:

"...Because a simple task could, theoretically be handled by a simple mechanism does not mean in fact that the brain handles it in that way. In an uncomplicated nervous system, yes; but in the complex brain of a higher animal, other mechanisms may insist in getting into the act and turn the simple task into a complex one... They exemplify an important principle, that the large brain, like large government may not be able to do simple things in a simple way."

(As part of his evidence, Hebb quotes the fact that a chimpanzee takes 200, 300 or even 400 trials to learn certain tasks that are learned by the rat in 10 or 20 trials.)

Two points with regard to computer simulation generally should perhaps be made: The first has been well stated by Tomkins (1963, Pp.9-10)

"... computer simulation must be judged, not by its resemblance to human beings, but by its conjoint economy, explanatory and predictive power."

The second is the more fundamental:

"... the objection has been made that symbols manipulated in simulation of personality may represent affective properties but are not in fact affects. This statement is certainly true, but /
but to frame it as an objection misses the point that computer simulations serve as models and that the distinction drawn between the representation of an object and the object itself is intrinsic to the function of models as representative systems" Messick (1963, Pp.311-312).

Finally in these suggestions for future research it might be possible to anticipate (and to try to answer) some of the more specific criticisms that might be made of this attempt and of Piaget's theory:

(i) It might be argued that Piaget's theory, impressive as it is, is not in fact sufficiently rigorous to generate predictions by the truth of which the theory would stand or fall. This may be so, but the question is an empirical one, and the easiest way to check it is surely by attempting a testing programme of the kind outlined here. Should it prove impossible to make adequate predictions from the theory as it stands, then we shall already have achieved information concerning weaknesses of Piaget's system, and may then be in a position to improve the system so that the testing programme can go ahead.

(ii) Our test situations must obviously be unambiguous (a characteristic not always true of Piaget's original experiments) and ones which are clearly related to the theory - i.e. it should not be possible to explain the same facts as well by other theories.

(iii) Since it might be argued that Piaget gets his results because he takes situations unfamiliar to the child (e.g. men and walking-sticks) we should ensure that we work within the framework of the child's own experience /
experience as much as possible, as Smirnov (1953), in an oblique reference, suggests.

(iv) Throughout the undertaking, conventional experimental controls must be observed: e.g. it is obviously necessary to check inter-examiner reliability in such operations as assigning behaviour-to-be-simulated to Piaget's theoretical categories.
1. **Comparison of the overlap and non-overlap forms of partially quantified problems, question by question**

The results quoted so far have been for children's total scores in a series of problems. In the case of the most important kind of problem — the partially quantified three-term series problem — these total scores were derived from children's attempts to solve eight problems, problems J to Q inclusive.

It is also possible to compare scores over children for matched questions. For example, we can compare the frequencies achieved by children for the overlap and non-overlap forms of question J and compare the direction of these scores (whether the overlap or the non-overlap form had the greater frequency) with the direction of the group results. In this way, it is possible to check whether each individual question pulls in the same way as the group results or whether these are questions which give results different from those for all questions combined.

This comparison has been made, using three different criteria: first of all, a comparison of the frequencies in the overlap and non-overlap forms of children who were correct in both of the questions which follow partially quantified problem-statements; secondly, a comparison of the frequencies of children making overlap error in the two forms; and finally a comparison of the frequencies of children who were correct on the second (and important) question following the problem-statement.

It /
It will be seen from the results that in all cases, all questions pull in the same way. There are no reversals.

**TABLE 19(a)**

*Partially quantified problems: correct responses in both questions*

<table>
<thead>
<tr>
<th>Age</th>
<th>overlap</th>
<th>non-overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6 1 8 4 5 3 5 4</td>
<td>1 0 0 0 1 0 0</td>
</tr>
<tr>
<td>9</td>
<td>6 5 7 6 9 6 8 5</td>
<td>4 2 2 1 3 2 4 2</td>
</tr>
<tr>
<td>10</td>
<td>11 5 10 8 8 9 10 9</td>
<td>3 1 3 2 4 4 4 3</td>
</tr>
</tbody>
</table>

**TABLE 19(b)**

*Partially quantified problems: correct responses in second question*

<table>
<thead>
<tr>
<th>Age</th>
<th>overlap</th>
<th>non-overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6 3 8 8 6 7 5 6</td>
<td>1 0 0 0 1 1 0 0</td>
</tr>
<tr>
<td>9</td>
<td>6 7 7 6 9 8 8 7</td>
<td>4 2 2 2 3 2 4 2</td>
</tr>
<tr>
<td>10</td>
<td>11 8 10 12 10 10 11</td>
<td>4 2 4 5 4 5 5 6</td>
</tr>
</tbody>
</table>

**TABLE 19(c)** /
**TABLE 19(a)**

Partially quantified problems: overlap error frequencies

<table>
<thead>
<tr>
<th>Age</th>
<th>overlap</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>non-overlap</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>overlap</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>non-overlap</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>overlap</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>non-overlap</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

No reversals
2. **Repeat of overlap/non-overlap study using fully quantified problems with asymmetry difficulties**

It might be argued (see page 78) that the difference between the statistical results of the fully-quantified problems (no significant difference between overlap and non-overlap problems) and the hypothesis derived from the clinical studies (that overlap problems are more difficult than non-overlap problems) could have been due to the difference in the problem-forms at the two stages. It will be recalled that the original clinical studies used overlap problems in which there were also asymmetry difficulties, whereas in the group studies, asymmetry difficulties were avoided. The avoidance of asymmetry difficulty necessitated a clumsy problem-statement in the non-overlap, but not in the overlap forms, and this might have accounted for the fact that the overlap problems were found to be no more difficult than the non-overlap in the group test.

Two new sets of matched problems were therefore constructed, one set of overlap problems, and one of non-overlap problems. These were identical with the problems used in the above group study, except for the fact that both types of problems involved asymmetry difficulty. Only fully-quantified forms were used, since the partially-quantified items used in the group study were not affected by the differential complexity of phrasing, and since the original (and divergent) clinical studies had been chiefly with fully quantified forms.

The other variables - relative size of first interval, type of relationship involved ('older than' or 'younger than'), sex of names used - were manipulated as before generating eight problems of each type (overlap and non-overlap) (see QA - QH on page 64). Spacers were interpolated as before.
Here, as an example, is Question C of this new asymmetry set, overlap version:

Donald has 9 marbles.

Donald has three marbles more than Ian. Donald has four marbles more than Brian.

So Ian has ___________ marbles.

So Brian has ___________ marbles.

The problems were administered as before to matched groups of children. The group sizes and ages were as for the previous occasion.

An analysis of variance was carried out, taking correct responses as criterion. Here are the results:

**TABLE 20(a)**

Overlap versus non-overlap problems: asymmetry forms

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap</td>
<td>( \xi x = 65 )</td>
<td>( \xi x = 57 )</td>
<td>( \xi x = 69 )</td>
</tr>
<tr>
<td>form</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
<tr>
<td>non-overlap</td>
<td>( \xi x = 32 )</td>
<td>( \xi x = 65 )</td>
<td>( \xi x = 74 )</td>
</tr>
<tr>
<td>form</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
<tr>
<td></td>
<td>( \xi x ) ( ov ) = 191</td>
<td>( \xi x ) ( n-ov ) = 171</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 20(b)**

\[ \begin{array}{c|c|c|c}
& age 8 & age 9 & age 10 \\
\hline
\text{overlap} & \xi x = 65 & \xi x = 57 & \xi x = 69 \\
\text{form} & (n = 13) & (n = 13) & (n = 13) \\
\hline
\text{non-overlap} & \xi x = 32 & \xi x = 65 & \xi x = 74 \\
\text{form} & (n = 13) & (n = 13) & (n = 13) \\
\hline
\xi x & 97 & 122 & 143 \\
\end{array} \]

1 For complete set of problems, see Appendix III, pages xxxvii ff.

2 For full calculation see Appendix II, page ix.
TABLE 20(b)

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form versus non-overlap form</td>
<td>1</td>
<td>5.128</td>
<td>5.128</td>
<td>less than one</td>
<td>-</td>
</tr>
<tr>
<td>Contrast among ages</td>
<td>2</td>
<td>40.795</td>
<td>20.398</td>
<td>3.710(2,72)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Forms X Ages (Interaction term)</td>
<td>2</td>
<td>40.180</td>
<td>20.090</td>
<td>3.654(2,72)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Within cells</td>
<td>72</td>
<td>395.846</td>
<td>5.498</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>481.949</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Over all three ages, then, the difference between the overlap asymmetry and the non-overlap asymmetry forms is not significant. At age eight, however, the pattern is quite different from that at age nine\(^1\) or age ten. There the frequency of correct responses in the overlap form is much greater than the frequency for the non-overlap form and the results are in the same direction as before. (This age-difference probably accounts for the significant interaction factor.)

A t-test was made of the difference at age eight, and this was found to be significant at the 5\(\%\) level\(^2\). However, if the reader will compare the present developmental pattern for overlap problems with that for similar problems in Group B (where larger numbers were used) he will see that the results for the eight-year-olds differ on the two occasions. Now Questions B12 /

---

\(^1\) The difference at age 9 was checked by t-test, and found to be non-significant.

\(^2\) For details of calculations see Appendix II, page x.
B12 and B14 in Group B also involved asymmetry difficulty, so it seems possible that the results in the present instance may be due to chance, in spite of the significance of the difference. If we can accept this explanation (and I can see no other) then the present results confirm those of the previous group test; if we do not accept it, then the present results go beyond the previous fully quantified results and support the partially quantified findings. In either case, we can say that the difference between results using the earlier (non-asymmetry) problem-forms, and the clinical expectations was not due to the fact that in clinical work we used chiefly the asymmetry forms, whereas in the group test we did not.

3. Repeat of the overlap/non-overlap study using partially quantified overlap problems with the link at the beginning

In the main design it was decided to make the partially quantified items as nearly parallel to the fully quantified items as possible. The fully quantified items were, of course, designed to avoid asymmetry difficulty (see page 64), and as a result were of the form 'A is 10. B is three years older than A. C is one year older than A.' Since the partially quantified problems were an exact parallel, they were also stated in the form 'B is three years older than A. C is one year older than A.' As a result, one possible interpretation of the finding that there is a significantly greater frequency of overlap problems correct than non-overlap is that children were using the process described on pages 47 f., in which they convert relatives to absolutes and treat 'B is three years older than A' as 'B is three years old', and 'C is one year older than A' as C is one year old'. In overlap problems, of course, such a process would lead to the correct answer /

\[1\] Of course there are differences in complexity between the two occasions, but this is unlikely to affect the issue.
answer. (In non-overlap problems, the process is irrelevant to solution, since A is three years older than B, B is one year older than C gives quantification only to one of the variables whose values the children are asked to find.) In case this were a frequent phenomenon, a parallel study was made, using partially quantified problems which gave no opportunity for this error. These were of the form 'A is three years older than B. A is one year older than C' and the process would thus lead to the obtaining of two different values for A. Using these problems¹ (which were otherwise identical with the previous partially quantified form) the results using correct responses as criterion were as follows:

TABLE 21(a)

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td>( \leq x = 48 )</td>
<td>( \leq x = 50 )</td>
<td>( \leq x = 56 )</td>
</tr>
<tr>
<td></td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
<tr>
<td>non-overlap form</td>
<td>( \leq x = 14 )</td>
<td>( \leq x = 15 )</td>
<td>( \leq x = 56 )</td>
</tr>
<tr>
<td></td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
</tbody>
</table>

The difference between the overlap and non-overlap² forms is thus in the same direction as before, and the difference is again significant at the 1% level³.

¹ For actual problem-forms, see Appendix III, pages xliiff.
² The non-overlap forms were identical with those used in the original study.
³ For details of calculation see Appendix II, page xi.
TABLE 21(b)

<table>
<thead>
<tr>
<th>Treatments</th>
<th>d.f.</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form V, non-overlap form</td>
<td>1</td>
<td>61.038</td>
<td>61.038</td>
<td>8.314(1,72)</td>
<td>&lt;01</td>
</tr>
<tr>
<td>Within cells</td>
<td>72</td>
<td>525.614</td>
<td>7.342</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results using correct response as criterion, are therefore generally in the same direction (using these problem-forms) as they were when the original problem-forms was used. And at ages 8 and 9, the numerical results on the two occasions are very close. On the present occasion, however, the results at age 10 present an anomaly. They neither fit the pattern of results for the younger ages on this occasion, nor do they agree with the results for age 10 on the previous occasion of testing (where, it will be recalled, the non-overlap problems were identical with those used here). Because of this, estimates of the children's ability were sought from the teacher, and when a list of these had been drawn up independently they were compared with the results for the present investigation. Now we would expect, both from the results from the younger children on this occasion and from the previous set of results, that the frequency of correct solutions for the overlap treatment would have exceeded the frequency for the non-overlap treatment, and that these results might be explained if we had the brighter children in the non-overlap group and the duller children in the overlap group. Although /
Although there is only one point difference in the average I.Q. for each group, four out of five of the class who, according to the teacher, were "working below their capacity" had been allocated to the overlap group, whereas all four of those who were "working above their measured I.Q." had been allocated to the non-overlap group. Now while this measure is admittedly crude and ex post facto it is in agreement with our prediction. (Confirmation will of course have to be sought by repeating the investigation with another group of 10-year-olds.)

The primary criterion used in the initial study was the occurrence of overlap error on each of the problem-forms, and the chief purpose of the present investigation was to discover whether the low frequency of overlap error in the original overlap forms was due to the 'loophole' through which children were able to solve these problems by converting relatives to absolutes. In fact the new problem-form gave frequencies of overlap error which were substantially similar to those in the original study:

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>original study</td>
<td>( \xi = 0 )</td>
<td>( \xi = 2 )</td>
<td>( \xi = 2 )</td>
</tr>
<tr>
<td>study</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
<tr>
<td>present study</td>
<td>( \xi = 1 )</td>
<td>( \xi = 5 )</td>
<td>( \xi = 2 )</td>
</tr>
<tr>
<td>study</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
<td>( n = 13 )</td>
</tr>
</tbody>
</table>

1 This has since been done with the result that the 10-year-olds fall into line with the results for the other ages. For the overlap form \( \xi x = 63 \) and for the non-overlap form \( \xi x = 49 \).
The frequencies are too small to compare by analysis of variance and since each cell total is derived from 104 instances (13 subjects tackling 8 problems each) the differences on the two occasions are almost certainly meaningless.

It therefore seems to be the case that the differences in frequencies of overlap error and of correct responses between overlap and non-overlap forms in the original study were not due to a loophole in the overlap form of the problem, and that the results therefore are valid.
4. Experimental design for projected attempt to discover the relative importance of 'return-to-link' and 'inclusion' for overlap error

During the clinical studies with overlap problems, it became apparent that overlap error could be the result of two difficulties - 'return to the link' and 'inclusion' (for explanation, see below). An experimental design was therefore evolved in an attempt to test the relative importance of these two processes for overlap error. Although (as a result of the statistical investigation) the present status of overlap error in overlap problems is uncertain (see discussion on pages 116f.), it might be worth summarising the design which was evolved, both because the orientation in the present report is largely historical and because these designs may take on new meaning once the clinical study of the significance of the statistical results is complete.

Once again, the kinds of problem-form used in the final design were evolved as a direct result of the solutions produced by the children to the original overlap problem. Now we named the overlap problem as we did because initially we saw in it the formal requirement that one relationship be included within the other. This requirement was subsequently named inclusion. The correct solution to the overlap problem may be represented thus:

(a)  
\[ \begin{array}{ccc} 
A & C & B \\
\end{array} \]

\[ \begin{array}{c} 
\text{e.g.} \{ A \text{ is three years older than } B \} \\
\{ A \text{ is one year older than } C. \} \\
\end{array} \]

and /.
and from this diagram the formal requirement of inclusion can be seen quite easily. In the same way, the 'classical' overlap error may be represented thus:

\[ (b) \]

\[ \begin{array}{c}
  \text{A} \\
  \text{B} \\
  \text{C}
\end{array} \]

(A is three years older than B  
\( \text{e.g.} \) B is one year older than C.

Here, obviously, there is no inclusion of one relationship within the other.

However, instead of the classical overlap error illustrated above, two other solutions to the overlap problem were frequently produced. (These are the solutions referred to above from which the final designs were evolved.) In the first of these, the child calculated the first interval correctly and then added to the value of the first dependent variable to achieve the value of the second. This procedure can be represented in the following diagram:

\[ (c) \]

\[ \begin{array}{c}
  \text{A} \\
  \text{C} \\
  \text{B}
\end{array} \]

Thirteen of the sixty-three subjects in the clinical sample from Group B made this error and these represented all age groups.

Mary gave a particularly clear example of this method of solution:

We /

---

1 This diagram fits equally well the case where the link changes value (see pages 164 ff.)
We want to find out the ages of Ian and David.

We know that Peter is 8, and that he is three years older than Ian and one year older than David. How old are Ian and David?

S: 'Ian's five. And David's six.

E: Now how did you know that?

S: I took three years off that boy's age and took one year on that boy.

E: I see. You added one and five to get the six?

S: Yes'.

In the second of the alternative solutions, the children made an asymmetry error on one of the relationships only:

(a) \[ \begin{array}{c} C \rightarrow A \leftarrow B \end{array} \]

This solution was given by five children of various ages. Here is Ann's answer:

We want to find out the ages of Jean and May.

We know that Betty is 10, and that she is one year older than May and three years older than Jean. How old are Jean and May?

S: 'If May is one year older than Betty, she's eleven. And if she's ...

Betty's three years older than Jean, Jean must be seven.

E: Now read it over to yourself again.

S: We know that Betty is ten ... (subject looks surprised.)

E: /
E: What is it?
S: That she is one year older than May. May's nine.
E: What age would Jean be then? Would Jean still be the same age?
S: No.
E: What age will she be?
S: (Pause) Six.
E: Why will she be six?
S: Because May is three years older than her. (Note how both variables were related directly to the link in the initial asymmetry interpretation, but how, when the asymmetry error was abandoned, the child resorted to classical overlap error. This strongly suggested that the original interpretation was indeed a form of overlap error and not simply asymmetry error on a single variable. Of the five children who made this error, four slipped into classical overlap error as soon as the asymmetry error was corrected.)

Solution (c) suggested that, as in the case of full overlap error - solution (b) - the difficulty for children in overlap problems lay in returning to the link. Unlike solution (b), however, solution (c) involved inclusion. Now if this inclusion was not merely accidental but was conceived by the child as such, this solution would also provide us with evidence (of a clinical sort) against inclusion as an overlap difficulty. (We could not, of course, be sure of this from the information given in the protocols alone. The problem-forms described below were, however, designed to check statistically whether this were so.)

Solution /
Solution (d), however, to some extent suggested the opposite interpretation: that the difficulty did not lie in the returning to the link which formally this solution demands, but in inclusion, which is formally absent from this solution. At the level of the children's thinking we could not be sure that they were in fact returning to the link. In Ann's protocol (quoted above), for example, she rather treated the relationships as a non-overlapping three-term series: May-Betty: Betty-Jean. In none of these protocols was there clear evidence of a return to the link. And the evidence on inclusion was here only the evidence of omission as against the more positive evidence provided by solution (c) on this point.

It could, however, have been the case that both return to the link and inclusion were experienced by the children as difficulties and that in solution (c) the one was being omitted and in solution (d) the other. The fact that solution (b) required neither would thus have accounted for its greater frequency.

What we had then was a hypothesis that overlap difficulty (in overlap problems) consists in failure to deal adequately with either or both of these formal requirements: inclusion, and the need to return to the link.

The alternative solutions to the overlap problem spontaneously produced by children in the 'thinking aloud' situation suggested problem-forms which might be used to check the relative importance of these formal requirements for the solution of the overlap problem:

The principle involved was to compare the results of children who were presented with a problem involving inclusion only (c.f. Fig. c) with the results of a matched group tackling problems involving return to the link only (c.f. Fig. d) and to compare each of these results with a matched group of /
of children tackling problems involving both of these difficulties - i.e. the classical 'overlap' problem (see Fig. a). In other words, in the classical 'overlap' problem there were thought to be two formal requirements: return to the link and inclusion. By presenting children with problems involving one of these supposed difficulties each, it was hoped to get an estimate of the relative psychological importance of these requirements for overlap difficulty.

In designing these items (and in refining the designs through clinical analysis) it was found to be impossible to generate items which were free of irrelevant interference factors: either asymmetry difficulties or change of direction difficulties.

Two designs were therefore attempted, either or both of which could be used. In each, an attempt was made to control the effect of one of these interference factors.

A. Avoiding asymmetry difficulties, but introducing change of direction difficulties

(In each design three problem forms were used, each in two versions: fully quantified and partially quantified.)

**Fully quantified forms**

( 1) **Overlap form** (formally requiring return to link and inclusion)

```
A   C   B
```

(A is 6
 B is 3 years younger than A; C is 1 year younger than A.
 (How old are B and C?)

For /

1 change of direction between problems, not necessarily within problems.
(The effect of change of direction could easily be tested by a separate design: a valuable future piece of research for its own sake.)
For each form, possible structural errors were worked out. Here, as an example, are the possible structural errors for the overlap form:

<table>
<thead>
<tr>
<th>Subjects' response</th>
<th>Possible significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>B = 3; C = 2</td>
<td>Return to link and/or inclusion as difficulties</td>
</tr>
<tr>
<td>B = 3; C = 4</td>
<td>Return to link only difficulty</td>
</tr>
<tr>
<td>B = 3; C = 7</td>
<td>Inclusion only difficulty</td>
</tr>
</tbody>
</table>

('Mirror-images' due to asymmetry error were also allowed for, but were considered to be unlikely.)

(ii) **Inclusion form**

(A is 6
B is 3 years older than A; C is 1 year younger than B.

**e.g.**

How old are B and C?

(iii) **Return to link form**

(A is 6
B is 3 years older than A; C is 1 year younger than A.

**e.g.**

How old are B and C?

Partially quantified forms were also constructed which were formally (and it was thought psychologically) parallel to these fully quantified forms:

**e.g.** /

---

1 Various other comparisons among problem-forms were also planned.
e.g. **Overlap form:**

(A is 6
B is 3 years older than A; A is 1 year younger than C.

B is 3 years younger than A. C is 1 year younger than A.

e.g.  
- Who is older, B or C?
- How much older is he?

B. **Avoiding change of direction difficulties,**

*but introducing asymmetry difficulties*

**Fully quantified forms**

(i) **Overlap form** (again formally requiring return to link and inclusion)

(A is 6
B is 3 years older than A; A is 1 year younger than C.

B is 3 years younger than A. C is 1 year younger than A.

(e.g. 
- Who is older, B or C?
- How much older is he?)

(ii) **Inclusion form:** identical with that in design A.

(iii) **Return to link form:** identical with that in design A.

Parallel partially quantified forms were also constructed in the same way as for design A.

It should perhaps be noted also that, clinically, overlap error seemed to be produced in at least two different ways. (When I speak in the past tense, I do not intend to convey that this clinical evidence is now considered invalid - see discussion on pages 116f.) Of these, the first was the more frequent. In this mode of solution, the child discovered the value of the second /
second dependent term by subtracting not from the link but from the first dependent term. This has already been illustrated by an excerpt from Jessie’s protocol (see pages 33 f.).

The second major process by which children arrived at overlap error is that in which the value of the link is changed when the child comes to deal with the second premiss. The protocol of Janice, quoted on pages 34 f. gave an example of this.

It is important to notice, however, that this discussion refers only to the problem-form where the link is mentioned once in the 'older than' version and twice in the 'younger than' version. It seems from the results that, where the link is not repeated, all or most of the children subtract from the dependent term; and where the link is repeated, the majority (but not all) of the children subtract from the link, but change its value. (Since the change in frequency of mention of the link is accompanied by a change in the direction of the statement of the asymmetrical relation, however, we cannot be sure that this result is final. I have not used the new form - in which the link is always repeated - sufficiently often at the clinical level to give any further evidence at the moment; and it is difficult to see how such information could be gained by group test.)

The results so far can be summed up roughly by the following table:

TABLE 23
TABLE 23

Frequency of occurrence of processes which lead to overlap error according to problem form

<table>
<thead>
<tr>
<th></th>
<th>Subtract from dependent variable</th>
<th>Change value of link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link not repeated</td>
<td>7x</td>
<td>-</td>
</tr>
<tr>
<td>Link repeated</td>
<td>x</td>
<td>4x</td>
</tr>
</tbody>
</table>

It can be seen that the new form in which the link is repeated does not guarantee uniform method of solution.

5. /

---

1 The results are given as multiples of x to emphasise the fact that we cannot rely on figures derived from a clinical procedure in which not all problems were given to all subjects.
5. Note on the assumptions underlying the estimate of atmosphere effect

In the discussion on pages 91 ff. two estimates were made of the incidence of atmosphere effect: one depended on the difference in incidence of correct responses - curves (E - F) - and the other on the difference in incidence of asymmetry error - curves (G - H). Now these two estimates are in fact derived from different sources: the former is an estimate of atmosphere effect in non-asymmetry problems, the latter in asymmetry problems. Since this is the case, assuming for the moment that atmosphere effect is in fact equal in the two problem-forms (but that it causes asymmetry error in one and coincides with the numerically-correct response in the other), what we would expect is a kind of 'sampling error'. This sampling error will be reflected in differential executive error, but should not cause executive errors in the two forms to differ significantly. In the same way, the validity of the two sources of estimate depends on different assumptions - the estimate derived from (E - F) assumes that the incidence of correct-responses in the two problem-forms would not otherwise differ significantly, also that what we might call 'coincident arbitrary error' (those arbitrary errors which coincide with the correct response in each case) behaves in the same way; whereas the (G - H) estimate assumes that 'non-coincident arbitrary error' (those which do not coincide with the correct response but with the asymmetry error value) would not otherwise differ significantly.¹ There is no reason to suppose that these conditions are not satisfied.

These distinctions underlie the discussion in the main text, and explain more fully the (non-significant) differences in asymmetry error in the two estimates.

¹ The (G - H) estimate of set effect is probably the better, since the frequency of arbitrary error, as estimated from curve H, is relatively low.
6. Discussion of the relation between atmosphere effect and asymmetry error

It might be argued that as a result of the research reported in this thesis, we should now dispense with the term 'asymmetry error' altogether. As a counsel for future usage this seems a reasonable course of action, and one which should be adopted: asymmetry error is, in the majority of cases, merely the product of atmosphere effect and a particular problem-form and, as we have seen, the atmosphere effect can occur in two-term series problems in the absence of asymmetry error.

In the description of the experimental investigation, however, I have found it useful to retain the concept of asymmetry error as a label for the kind of process I expected to occur in certain problems if in fact atmosphere effect were operating. Without the use of this label, description would have become very much more difficult. For example, on page 88 we have the statement: "The numbers of asymmetry errors occurring at each age... are given in the following table". Without the label 'asymmetry errors' or an equivalent (and since 'asymmetry error' has historical connections with Donaldson's work it seemed as good as any) we would have been forced to say something like this: "The numbers of wrong solutions of the kind we would expect to occur if in fact atmosphere effect is operating, in those problems which predispose to wrong solution under these conditions, together with the equivalent wrong solutions in the problems in which we expect atmosphere effect to operate but not to produce wrong solution, are given in the following table".

1 Where it is not, it can again be called by the name of the error to which it is related - e.g. arbitrary error.
APPENDIX II

Analysis of Variance Calculations
and Data from Longitudinal Study
TABLE 4 (a) and (b) (pages 68 - 69)

(a) data: frequencies of overlap error per subject (n = 13)

**Overlap form**
- **Age 8:** none
- **Age 9:** one child with two errors
- **Age 10:** " "

**Non-overlap form**
- **Age 8:** 2, 0, 2, 0, 7, 0, 7, 1, 0, 0, 0, 0, 0, 0, 0 = 20
- **Age 9:** 2, 1, 0, 2, 0, 0, 5, 0, 1, 0, 0, 0, 0, 0 = 11
- **Age 10:** 0, 0, 1, 0, 1, 4, 1, 0, 0, 1, 7, 0, 5 = 20

(b) calculations

<table>
<thead>
<tr>
<th></th>
<th>Age 8</th>
<th>Age 9</th>
<th>Age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overlap form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sum x) = 0</td>
<td>(\sum x) = 2</td>
<td>(\sum x) = 2</td>
<td></td>
</tr>
<tr>
<td>(\sum x^2) = 0</td>
<td>(\sum x^2) = 4</td>
<td>(\sum x^2) = 4</td>
<td></td>
</tr>
<tr>
<td>(\sum x^2) = 0</td>
<td>(\sum x^2) = 3,692</td>
<td>(\sum x^2) = 3,692</td>
<td></td>
</tr>
<tr>
<td><strong>Non-overlap form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sum x) = 20</td>
<td>(\sum x) = 11</td>
<td>(\sum x) = 20</td>
<td></td>
</tr>
<tr>
<td>(\sum x^2) = 108</td>
<td>(\sum x^2) = 35</td>
<td>(\sum x^2) = 94</td>
<td></td>
</tr>
<tr>
<td>(\sum x^2) = 77,231</td>
<td>(\sum x^2) = 25,926</td>
<td>(\sum x^2) = 63,231</td>
<td></td>
</tr>
</tbody>
</table>

\(\sum x_{ov.} = 4\)
\(\sum x_{n.ov.} = 51\)
\(G.T. = 55\)

**Sums of Squares: Forms**

\(S = \frac{47^2}{78} = 2209 = 28.321\)

**Sums of Squares: Ages**

\(S = \frac{1}{26} \times \frac{1053}{1} = \frac{55^2}{78}\)

\(= 1053 - 3025\)

\(= 40,500 - 38,782\)

\(= 1,718\)

**Sums of Squares: Interaction Term**

\(S = \frac{1}{26} \times \frac{805}{1} = 28.321\)

\(= 50,962 - 28,321\)

\(= 2,641\)

**Sums of Squares Within Cells**

\(= 173,538\)

**Total Sums of Squares**

\(= 245 - 38.782\)

\(= 206,218\)
### TABLE 5 (a) and (b) (pages 71 - 72)

(a) **Data:** frequencies of failure-to-respond per subject \( (n = 13) \)

#### Overlap Form

<table>
<thead>
<tr>
<th>Age</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\( \sum x = 34 \)

\[ \sum x^2 = 192 \]

\[ \chi^2 = 103.077 \]

#### Non-Overlap Form

<table>
<thead>
<tr>
<th>Age</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\( \sum x = 27 \)

\[ \sum x^2 = 341 \]

\[ \chi^2 = 84.923 \]

#### Calculations

<table>
<thead>
<tr>
<th>Age</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sum x )</td>
<td>34</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>( \sum x^2 )</td>
<td>192</td>
<td>141</td>
<td>1</td>
</tr>
</tbody>
</table>

\( \sum x_{ov.} = 62 \)

<table>
<thead>
<tr>
<th>Age</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sum x )</td>
<td>50</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>( \sum x^2 )</td>
<td>298</td>
<td>250</td>
<td>2</td>
</tr>
</tbody>
</table>

\( \sum x_{n.ov.} = 92 \)

#### Sums of Squares: Forms

\[
\sum x = 30^2 \quad = \quad 900 \quad = \quad 11.538
\]

\[
\sum x^2 = 78 \quad \quad = \quad 78 \quad \quad = \quad 11.538
\]

#### Sums of Squares: Ages

\[
\sum x = 11554 \quad - \quad (154)^2
\]

\[
\sum x^2 = 23716 \quad = \quad 11554 \quad - \quad 26 \quad \quad = \quad 23716 \quad \quad = \quad 11554 \quad - \quad 26 \quad \quad = \quad 23716
\]

\[
= \quad 444.385 \quad - \quad 304.051 \quad = \quad 140.334
\]

#### Sums of Squares: Interaction Term

\[
\sum x = 4.26 \quad - \quad 11.538
\]

\[
\sum x^2 = 16.385 \quad - \quad 11.538 \quad = \quad 4.847
\]

#### Sums of Squares Within Cells

\[
= \quad 423.230
\]

#### Total Sums of Squares

\[
= \quad 884 \quad - \quad 304.051
\]

\[
= \quad 579.949
\]
TABLE 6 (a) and (b) (page 73)

(a) data: frequencies of use of interval-as-answer per subject (n = 13)

<table>
<thead>
<tr>
<th>Age 8</th>
<th>Age 9</th>
<th>Age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0, 5, 1, 2, 0, 1, 0, 0, 0, 8, 0, 5</td>
<td>0, 1, 0, 3, 0, 2, 2, 0, 0, 2, 1</td>
<td>0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0</td>
</tr>
<tr>
<td>Non-overlap form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3, 0, 1, 3, 0, 8, 1, 3, 2, 0, 5, 0</td>
<td>2, 4, 0, 6, 1, 0, 0, 1, 2, 1, 8</td>
<td>0, 2, 1, 4, 0, 3, 4, 0, 7, 1, 8, 1</td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th>Age 8</th>
<th>Age 9</th>
<th>Age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{x} = 22$</td>
<td>$\bar{x} = 18$</td>
<td>$\bar{x} = 5$</td>
</tr>
<tr>
<td>$\bar{x}^2 = 120$</td>
<td>$\bar{x}^2 = 72$</td>
<td>$\bar{x}^2 = 9$</td>
</tr>
<tr>
<td>$\bar{x}'^2 = 82.769$</td>
<td>$\bar{x}'^2 = 47.077$</td>
<td>$\bar{x}'^2 = 7.077$</td>
</tr>
<tr>
<td>Non-overlap form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{x} = 26$</td>
<td>$\bar{x} = 25$</td>
<td>$\bar{x} = 31$</td>
</tr>
<tr>
<td>$\bar{x}^2 = 122$</td>
<td>$\bar{x}^2 = 127$</td>
<td>$\bar{x}^2 = 161$</td>
</tr>
<tr>
<td>$\bar{x}'^2 = 70.000$</td>
<td>$\bar{x}'^2 = 78.923$</td>
<td>$\bar{x}'^2 = 87.077$</td>
</tr>
</tbody>
</table>

Sums of Squares: Forms = $\frac{37^2}{78} - \frac{1369}{78} = 17.551$

Sums of Squares: Ages = $\frac{5449 - 127^2}{26} = 209.577 - 206.782 = 2.795$

Sums of Squares: Interaction Term = $\frac{741}{26} - 17.551$

Sums of Squares Within Cells = 372.923

Total Sums of Squares = 611 - 206.782 = 404.218
TABLE 8 (a) and (b) (pages 75-76.)

(a) data: frequencies of correct response to first question, per subject
(n = 13)

overlap form

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>4</th>
<th>0</th>
<th>4</th>
<th>0</th>
<th>5</th>
<th>2</th>
<th>5</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>7</th>
<th>1</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>age 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

non-overlap form

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>1</th>
<th>3</th>
<th>2</th>
<th>4</th>
<th>0</th>
<th>5</th>
<th>1</th>
<th>1</th>
<th>7</th>
<th>2</th>
<th>3</th>
<th></th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>age 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x = 51$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 281$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x'^2 = 80.923$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x = 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 155$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x'^2 = 60.769$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 412$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x'^2 = 69.231$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 122$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x'^2 = 53.231$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 86$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 135$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 51$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 529$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 58$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 312$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 64$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum x^2 = 77$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sums of Squares: Forms = $\frac{412}{78} = 1681/78 = 21.551$

Sums of Squares: Ages = $40505 - \frac{343^2}{26} = 1557.885 - 1508.320 = 49.565$

Sums of Squares: Interaction Term = $\frac{653 - 21.551}{26} = 25.115 - 21.551 = 3.564$

Sums of Squares Within Cells = $424,000$

Total Sums of Squares = $2007,000 - 1508.320 = 498,680$
TABLE 9 (a) and (b) (page 77.)

(a) data: frequencies of correct responses in fully quantified problems, per subject

<table>
<thead>
<tr>
<th>Age</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>4</th>
<th>0</th>
<th>0</th>
<th>8</th>
<th>7</th>
<th>0</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>4</th>
<th>6</th>
<th>6</th>
<th>7</th>
<th>4</th>
<th>0</th>
<th>8</th>
<th>1</th>
<th>0</th>
<th>8</th>
<th>2</th>
<th>5</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th>Age</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90</td>
<td>152</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>198</td>
<td>189</td>
<td>G.T.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Age 8</th>
<th>Age 9</th>
<th>Age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td>(\sum x = 39)</td>
<td>(\sum x = 89)</td>
<td>(\sum x = 70)</td>
</tr>
<tr>
<td></td>
<td>(\sum x^2 = 259)</td>
<td>(\sum x^2 = 667)</td>
<td>(\sum x^2 = 510)</td>
</tr>
<tr>
<td></td>
<td>(\sum' x^2 = 142,000)</td>
<td>(\sum' x^2 = 57,692)</td>
<td>(\sum' x^2 = 133,077)</td>
</tr>
<tr>
<td>non-overlap form</td>
<td>(\sum y = 51)</td>
<td>(\sum y = 63)</td>
<td>(\sum y = 75)</td>
</tr>
<tr>
<td></td>
<td>(\sum y^2 = 111,923)</td>
<td>(\sum y^2 = 79,692)</td>
<td>(\sum y^2 = 86,308)</td>
</tr>
<tr>
<td></td>
<td>(\sum x y = 90)</td>
<td>(\sum x y = 152)</td>
<td>(\sum x y = 145)</td>
</tr>
</tbody>
</table>

Sums of Squares: Forms = \(\frac{9^2}{78} = 81\) = 1,039

Sums of Squares: Ages = \(\frac{52229 - 387^2}{26} = \frac{149769}{78} = 2008.808 - 1920.115 = 88.693\)

Sums of Squares: Interaction Term = \(\frac{84.5 - 1.039}{26} = \frac{32.500 - 1.039}{31.461}\)

Sums of Squares Within Cells = 609.692

Total Sums of Squares = 2651,000 - 1920.115 = 730,885
**TABLE 12 (a) and (b) (page 89)**

(a) data: frequencies per subject of asymmetry error

<table>
<thead>
<tr>
<th>asymmetry form</th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0, 0, 6, 11, 3, 2, 1, 2, 9, 7, 1, 1, 1</td>
<td>0, 0, 6, 10, 10, 3, 0, 3, 5, 1, 7, 7, 1</td>
<td>2, 2, 0, 5, 0, 4, 1, 1, 15, 1, 4, 1, 6</td>
<td>1, 10, 1, 8, 1, 3, 4, 2, 1, 6, 3, 4, 1, 2</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>62</td>
<td>46</td>
<td>44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>non-asymmetry form</th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1, 0, 1, 1, 3, 1, 1, 0, 2, 0, 0, 2, 0, 0</td>
<td>0, 3, 0, 2, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0</td>
<td>0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0</td>
<td>0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>8</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
<th>( \Sigma x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry form</td>
<td>( \Sigma x ) = 44</td>
<td>( \Sigma x ) = 62</td>
<td>( \Sigma x ) = 46</td>
<td>( \Sigma x ) = 44</td>
</tr>
<tr>
<td>form</td>
<td>( \Sigma x^2 ) = 308</td>
<td>( \Sigma x^2 ) = 424</td>
<td>( \Sigma x^2 ) = 354</td>
<td>( \Sigma x^2 ) = 254</td>
</tr>
<tr>
<td></td>
<td>( \Sigma x^2 ) = 159,077</td>
<td>( \Sigma x^2 ) = 128,308</td>
<td>( \Sigma x^2 ) = 191,231</td>
<td>( \Sigma x^2 ) = 105,077</td>
</tr>
<tr>
<td>non-asymmetry form</td>
<td>( \Sigma x ) = 12</td>
<td>( \Sigma x ) = 8</td>
<td>( \Sigma x ) = 3</td>
<td>( \Sigma x ) = 6</td>
</tr>
<tr>
<td>form</td>
<td>( \Sigma x^2 ) = 22</td>
<td>( \Sigma x^2 ) = 16</td>
<td>( \Sigma x^2 ) = 5</td>
<td>( \Sigma x^2 ) = 14</td>
</tr>
<tr>
<td></td>
<td>( \Sigma x^2 ) = 10,923</td>
<td>( \Sigma x^2 ) = 11,077</td>
<td>( \Sigma x^2 ) = 4,308</td>
<td>( \Sigma x^2 ) = 11,231</td>
</tr>
</tbody>
</table>

\[ \sum_{x} = 196 \]

Sums of Squares: Forms: \( \frac{(167)^2}{104} = 278.89 = 268.163 \)

Sums of Squares: Ages: \( \frac{1 (3136 + 4900 + 2401 + 2500) - 486.779}{26} \)
\[ = \frac{497.577 - 486.779}{10.798} = 10.798 \]

Sums of Squares: Interaction Term: \( \frac{7233 - 268.163}{26} \)
\[ = \frac{278.192 - 268.163}{10.029} \]

Sums of Squares Within Cells \( = 621.232 \)

Total Sums of Squares \( = 1397 - 486.779 \)
\[ = 910.221 \]
(a) data: frequencies per subject of correct responses in two-term series problems

<table>
<thead>
<tr>
<th>Age</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 8:</td>
<td>19, 20, 8, 6, 10, 2, 2, 9, 6, 8, 19, 16, 0</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 9:</td>
<td>12, 14, 13, 4, 8, 16, 19, 16, 13, 19, 12, 13, 2</td>
<td>161</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 10:</td>
<td>11, 18, 18, 15, 15, 19, 17, 18, 18, 19, 12, 14, 19, 18</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 11:</td>
<td>19, 18, 19, 10, 19, 17, 18, 18, 19, 12, 14, 19, 18</td>
<td>208</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th>Age</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetry</td>
<td>125</td>
<td>161</td>
<td>200</td>
<td>208</td>
</tr>
<tr>
<td>Form</td>
<td>694</td>
<td>309</td>
<td>334</td>
<td>3502</td>
</tr>
<tr>
<td>Non-asymmetry</td>
<td>565.077</td>
<td>315.077</td>
<td>237.077</td>
<td>174.000</td>
</tr>
<tr>
<td>Form</td>
<td>219</td>
<td>219</td>
<td>219</td>
<td>219</td>
</tr>
<tr>
<td>Non-asymmetry</td>
<td>2726</td>
<td>4073</td>
<td>4758</td>
<td>502.923</td>
</tr>
<tr>
<td>Form</td>
<td>502.923</td>
<td>383.692</td>
<td>22.000</td>
<td>22.000</td>
</tr>
</tbody>
</table>

Sums of Squares: Form: \((190)^2 = 36100 = 34.712\)

Sums of Squares: Age: \(\frac{639154}{26} - \frac{(1578)^2}{104} = 639.73\)

Sums of Squares: Interaction Term: \(\frac{9214}{26} - \frac{34.712}{26} = 7.27\)

Sums of Squares Within Cells: \(2226.77\)

Total Sums of Squares: \(27164 - 23943.12 = 3220.89\)

TABLE 13 (a) and (b) (page 92)
TABLE 16 (a) and (b) (Pages 95 - 96)

(a) data: frequencies per subject of executive error in two-term series problems

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 8</td>
<td>1, 02, 0, 2, 7, 6, 1, 1, 3, 0, 0, 16 = 39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>2, 3, 0, 5, 0, 1, 1, 1, 2, 0, 0, 16 = 31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>5, 0, 1, 0, 1, 1, 0, 2, 0, 0, 0, 12 = 12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 11</td>
<td>0, 2, 0, 2, 0, 0, 0, 0, 2, 2, 0, 0, 8 = 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-asymmetry form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 8</td>
<td>2, 0, 0, 1, 0, 4, 0, 5, 2, 0, 12, 1, 0 = 27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>0, 2, 2, 0, 2, 0, 0, 0, 0, 0, 0, 18 = 32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>0, 0, 0, 0, 1, 0, 1, 5, 0, 0, 0, 8 = 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 11</td>
<td>0, 2, 1, 0, 2, 0, 1, 0, 0, 0, 0, 6 = 6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
<th>age 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>asymmetry</td>
<td>$\xi = 39$</td>
<td>$\xi = 31$</td>
<td>$\xi = 12$</td>
<td>$\xi = 8$</td>
</tr>
<tr>
<td>form</td>
<td>$\xi^2 = 361$</td>
<td>$\xi^2 = 301$</td>
<td>$\xi^2 = 36$</td>
<td>$\xi^2 = 16$</td>
</tr>
<tr>
<td></td>
<td>$\xi^2 = 244,000$</td>
<td>$\xi^2 = 227,077$</td>
<td>$\xi^2 = 24,923$</td>
<td>$\xi^2 = 11,077$</td>
</tr>
<tr>
<td>non-</td>
<td>$\xi = 27$</td>
<td>$\xi = 32$</td>
<td>$\xi = 8$</td>
<td>$\xi = 6$</td>
</tr>
<tr>
<td>asymmetry</td>
<td>$\xi^2 = 195$</td>
<td>$\xi^2 = 400$</td>
<td>$\xi^2 = 28$</td>
<td>$\xi^2 = 10$</td>
</tr>
<tr>
<td>form</td>
<td>$\xi^2 = 138,923$</td>
<td>$\xi^2 = 321,231$</td>
<td>$\xi^2 = 23,077$</td>
<td>$\xi^2 = 7,231$</td>
</tr>
<tr>
<td></td>
<td>$\sum \xi^2_8 = 66$</td>
<td>$\sum \xi^2_9 = 63$</td>
<td>$\sum \xi^2_{10} = 20$</td>
<td>$\sum \xi^2_{11} = 14$</td>
</tr>
</tbody>
</table>

Sums of Squares: Forms: $\frac{17^2}{104} = \frac{289}{104} = 2.779$

Sums of Squares: Ages: $\frac{8921 - 26569}{26} = \frac{343.115 - 255.471}{104} = 87.644$

Sums of Squares: Interaction Term: $\frac{165}{26} - 2.779 = 6.346 - 2.779 = 3.567$

Sums of Squares Within Cells: 997.539

Total Sums of Squares: 1091.529
TABLE 20 (a) and (b) (pages 150 - 151)

(a) data: frequencies per subject of correct responses in asymmetry forms of overlap/non-overlap problems

overlap forms with asymmetry

<table>
<thead>
<tr>
<th>Age</th>
<th>3, 8, 5, 8, 1, 0, 6, 3, 4, 6, 8, 5, 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 9</td>
<td>2, 4, 5, 7, 4, 0, 6, 6, 8, 8, 1, 2</td>
</tr>
<tr>
<td>Age 10</td>
<td>2, 4, 5, 7, 7, 6, 0, 5, 7, 7, 8</td>
</tr>
</tbody>
</table>

non-overlap form with asymmetry

| Age 8 | 2, 0, 4, 0, 1, 3, 0, 4, 5, 6 |
| Age 9 | 7, 3, 8, 7, 0, 4, 6, 5, 8, 5 |
| Age 10| 5, 8, 7, 6, 7, 1, 4, 5, 6, 5 |

(b) calculations

<table>
<thead>
<tr>
<th>Age 8</th>
<th>3, 8, 5, 8, 1, 0, 6, 3, 4, 6, 8, 5, 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 9</td>
<td>2, 4, 5, 7, 4, 0, 6, 6, 8, 8, 1, 2</td>
</tr>
<tr>
<td>Age 10</td>
<td>2, 4, 5, 7, 7, 6, 0, 5, 7, 7, 8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overlap Form</th>
<th>65</th>
<th>57</th>
<th>69</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Squares</td>
<td>413</td>
<td>331</td>
<td>439</td>
</tr>
<tr>
<td>Overlap Form</td>
<td>88,000</td>
<td>81,077</td>
<td>72,769</td>
</tr>
<tr>
<td>Sum of Squares: Non-Overlap Form</td>
<td>32</td>
<td>65</td>
<td>74</td>
</tr>
<tr>
<td>Sum of Squares</td>
<td>132</td>
<td>387</td>
<td>460</td>
</tr>
<tr>
<td>Sum of Squares: Interaction Term</td>
<td>53,231</td>
<td>62,000</td>
<td>38,769</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>97</th>
<th>122</th>
<th>143</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.T.</td>
<td>362</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sums of Squares: Forms = \( \frac{(20)^2}{78} = \frac{400}{78} = 5.128 \)

Sums of Squares: Ages = \( \frac{44742}{26} = \frac{131044}{78} = 1720.846 - 1680.051 = 40.795 \)

Sums of Squares: Interaction Term = \( \frac{1178}{26} - 5.128 \)

= \( 45.308 - 5.128 = 40.180 \)

Sums of Squares Within Cells = 395.846

Total Sum of Squares = 2162 - 1680.051

= 181.949
t-test of difference between means at age 8:

asymmetry forms of overlap/non-overlap problems

(see page 51)

\[ t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2}\right) \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \]

\[ = \frac{1}{13} \left(\frac{65 - 32}{\sqrt{\left(\frac{88,000 + 55,231}{24}\right) \left(\frac{2}{\frac{24}{13}}\right)}}\right) \]

\[ = \frac{33}{13} \left(\frac{141,231 \times 2^{\frac{2}{13}}}{\frac{24}{13}}\right) \]

\[ = 2.538 \left(\frac{5.685 \times 0.154}{\sqrt{0.906}}\right) \]

\[ = \frac{2.538}{0.952} \]

\[ = 2.666 \text{ (d.f. } = 24) \]

Significant beyond the 5% level
TABLE 21 (a) and (b) (pages 153 - 154)

(a) data: frequencies per subject of correct responses in three-term series problems where link is given first in the overlap form

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 8</td>
<td>0, 5, 7, 4, 2, 2, 6, 5, 7, 3, 0, 5, 2 = 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>7, 5, 5, 8, 2, 1, 0, 0, 7, 6, 2, 5, 2 = 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>1, 7, 4, 0, 7, 5, 0, 8, 6, 4, 8, 0, 6 = 56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-overlap form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 8</td>
<td>0, 1, 0, 2, 1, 0, 1, 0, 0, 1, 0, 7, 1 = 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 9</td>
<td>0, 0, 0, 8, 1, 5, 0, 0, 0, 0, 1 = 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 10</td>
<td>6, 0, 8, 0, 5, 7, 0, 2, 0, 6, 8, 7, 7 = 56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) calculations

<table>
<thead>
<tr>
<th></th>
<th>age 8</th>
<th>age 9</th>
<th>age 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>overlap form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Sigma x = 48)</td>
<td>(\Sigma x = 50)</td>
<td>(\Sigma x = 56)</td>
<td>(\sum x_{ov} = 154)</td>
</tr>
<tr>
<td>(\Sigma x^2 = 246)</td>
<td>(\Sigma x^2 = 286)</td>
<td>(\Sigma x^2 = 356)</td>
<td></td>
</tr>
<tr>
<td>(\Sigma'x^2 = 68.769)</td>
<td>(\Sigma'x^2 = 93.692)</td>
<td>(\Sigma'x^2 = 114.769)</td>
<td></td>
</tr>
<tr>
<td>non-overlap form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Sigma x = 14)</td>
<td>(\Sigma x = 15)</td>
<td>(\Sigma x = 56)</td>
<td>(\sum x_{n,ov} = 85)</td>
</tr>
<tr>
<td>(\Sigma x^2 = 58)</td>
<td>(\Sigma x^2 = 91)</td>
<td>(\Sigma x^2 = 376)</td>
<td></td>
</tr>
<tr>
<td>(\Sigma'x^2 = 42.923)</td>
<td>(\Sigma'x^2 = 73.692)</td>
<td>(\Sigma'x^2 = 134.769)</td>
<td></td>
</tr>
<tr>
<td>(\sum x_8 = 62)</td>
<td>(\sum x_9 = 65)</td>
<td>(\sum x_{10} = 112)</td>
<td>G.T. = 239</td>
</tr>
</tbody>
</table>

Sums of Squares: Forms: \(\frac{(69)^2}{78} = \frac{4761}{78} = 61.038\)

Sums of Squares Within Cells: 528.614
A. Overlap error data: numbers of overlap errors made by each child on the two occasions

(a) age 8 becoming age 9
   occ. 1  0 7 1 2 0 2 0 7 0 1 0  
   occ. 2  0 8 0 0 1 0 1 6 4 0 7  

(b) age 9 becoming age 10
   occ. 1  0 0 0 5 2 1 0 2 1 0 0  
   occ. 2  4 6 3 0 0 3 1 1 1 0 6  

(c) age 10 becoming age 11
   occ. 1  0 1 0 1 1 0 0 1 0 4 7 5  
   occ. 2  5 7 0 8 3 0 2 5 5 7 8 1  

B. Asymmetry error data: numbers of asymmetry errors made by each child on the two occasions

(a) age 8 becoming age 9
   occ. 1  7 1 1 1 6 1 2 1 2 0 0  
   occ. 2  2 3 3 9 4 7 0 8 2 4  

(b) age 9 becoming age 10
   occ. 1  5 7 7 6 3 0 1 0 1 0 6 1 3 1  
   occ. 2  2 0 3 4 1 1 6 0 0 0 0 0 0  

(c) age 10 becoming age 11
   occ. 1  6 1 1 4 4 5 0 2 0 2  
   occ. 2  2 1 2 2 1 1 0 1 1  

(d) age 11 becoming age 12
   occ. 1  8 2 1 4 1 1 0 4 3 6 1 1  
   occ. 2  4 1 0 2 0 6 0 2 1 1 1  

xii
APPENDIX III

Record of problems used in the group tests
Problems used in Group B

B12

We want to find out the ages of Jean and May. We know that Betty is 10, and that she is one year older than May, and three years older than Jean.

So Jean is ....... years old.
So May is ....... years old.

Sp.

A man walks along a straight road for 7 miles. Then he turns round and walks back along the same straight road for 3 miles.

So he is ....... miles from his starting point.

B15

We want to find out the ages of Bill and Harry.

We know that Jim is 7, and that Bill is three years younger than Jim, and that Harry is two years younger than Jim.

So Bill is ....... years old.
So Harry is ....... years old.

Sp.

My brother has one brother. He has no sisters. How many brothers have I?

Write your answer here: ......
We want to find out the ages of Ian and David. We know that Peter is 8, and that he is three years older than Ian, and one year older than David.

So Ian is \( \ldots \ldots \) years old.

So David is \( \ldots \ldots \) years old.

Two men start to walk along the same road in the same direction. One man walks 2 miles, and the other man walks 5 miles.

So they are \( \ldots \ldots \) miles apart at the end of their walk.

We want to find out the ages of Anne and Susan. We know that Joan is 9, and that Susan is two years younger than Joan, and that Anne is three years younger than Joan.

So Anne is \( \ldots \ldots \) years old.

So Susan is \( \ldots \ldots \) years old.

How many triangles are there in this picture?

Write your answer here: \( \ldots \ldots \)
Problems used in investigation into overlap error

0/A* Jean is 6.
Sally is three years older than Jean. Linda is one year older than Jean.
So Sally is ________ years old.
So Linda is ________ years old.

0/Sp.* Write sixteen in figures. ________

0/B A can holds 6 pints of milk.
A red jug holds three pints less than the can. A blue jug holds two pints less than the can.
So the red jug holds ________ pints.
So the blue jug holds ________ pints.

0/Sp. Multiply 7 and 8 ________

0/C Donald has 6 marbles.
Ian has two marbles more than Donald. Brian has five marbles more than Donald.
So Ian has ________ marbles.
So Brian has ________ marbles.

* Where 0/A = overlap form/question A and Sp. = spacer problem
Add 12 and 7

Joyce jumps 6 feet.
Ann jumps one foot less than Joyce. Nancy jumps three feet less than Joyce.
So Ann jumps ______ feet.
So Nancy jumps ______ feet.

One toy cost 6 shillings and another cost 7 shillings.
How many shillings did the two cost altogether? ______ shillings.

Colin has 5 shillings.
Bill has four shillings more than Colin. Eric has three shillings more than Colin.
So Bill has ______ shillings.
So Eric has ______ shillings.

Subtract 11 from 15

Alice has three books, one red, one green, and one blue.
The red book has 80 pages.
The green book has twenty pages less than the red book.
The blue book has ten pages less than the red book.
So the green book has _______ pages.
So the blue book has _______ pages.
Divide 15 by 5

Alan is 5.
Tony is one year older than Alan. Bob is four years older than Alan.
So Tony is ________years old.
So Bob is ________years old.

The houses on Bank Street have odd numbers on one side and even numbers on the other side. Tom Smith lives at number 27. What are the numbers on each side of Tom's house? ________and_______

Ian is 6 feet tall.
David is one foot smaller than Ian. Bruce is two feet smaller than Ian.
So David is ________feet tall.
So Bruce is ________feet tall.

Add 10 to 35

Doris is five years older than Grace. Edith is three years older than Grace.
Who is older, Doris or Edith? ________
How much older is she? ________years.
Multiply 5 and 8

Sandy has three jugs, one yellow, one green, and one white.
The yellow jug holds five pints less than the white one.
The green jug holds four pints less than the white one.
Which holds more, the yellow jug or the green jug?
How much more does it hold?

Write nineteen in figures

Joy has seven pennies more than Olive. Bess has eight pennies more than Olive.
Who has more, Joy or Bess?
How many more has she?

A dozen eggs cost 4 shillings. How much do 2 dozen cost?

Archie jumps two feet less than Dennis. Peter jumps three feet less than Dennis.
Who jumps further, Archie or Peter?
How much further does he jump?
Take 7 away from 19 ________

June has seven shillings more than Carol. Audrey has five shillings more than Carol.

Who has more, June or Audrey? ________

How much more has she? ________ shillings

Divide 9 by 3 ________

Joe has three books, one yellow, one black and one orange.

The yellow book has ten pages less than the orange book.
The black book has five pages less than the orange book.

Which has more pages, the yellow book or the black book? ________

How many more pages has it? ________ pages.

How many feet do four dogs have? ________ feet.

Cathy is five years older than Agnes. Jess is eight years older than Agnes.

Who is older, Cathy or Jess? ________

How much older is she? ________ years.

Add 4 to 9 ________
Bert has three marbles less than Charles. Jerry has five marbles less than Charles.

Who has more, Bert or Jerry? ________

How many more has he? ________ marbles.

If a dozen eggs cost 4 shillings, how much does one egg cost?

_______ pennies.

Jean is 6.

Sally is three years older than Jean. Linda is one year older than Sally.

So Sally is ________ years old.

So Linda is ________ years old.

Write sixteen in figures ________

A can holds 6 pints of milk.

A red jug holds three pints less than the can.

A blue jug holds two pints less than the red jug.

So the red jug holds ________ pints.

So the blue jug holds ________ pints.

Multiply 7 and 8 ________

* Where NO/A means non-overlap form, Question A
Donald has 6 marbles.  
Ian has two marbles more than Donald.  Brian has five marbles more than Ian.  
So Ian has ________ marbles.  
So Brian has ________ marbles.

Add 12 and 7 ________

Joyce jumps 6 feet.  
Ann jumps one foot less than Joyce.  Nancy jumps three feet less than Ann.  
So Ann jumps ________ feet.  
So Nancy jumps ________ feet.

One toy cost 6 shillings and another cost 7 shillings.  
How many shillings did the two cost altogether? ________ shillings.

Colin has 5 shillings.  
Bill has four shillings more than Colin.  Eric has three shillings more than Bill.  
So Bill has ________ shillings.  
So Eric has ________ shillings.

Subtract 11 from 15 ________
Alice has three books, one red, one green, and one blue.
The red book has 80 pages.
The green book has twenty pages less than the red book. The blue book has ten pages less than the green book.
So the green book has \_
So the blue book has \_\_

Divide 15 by 5 \_

Alan is 5.
Tony is one year older than Alan. Bob is four years older than Tony.
So Tony is \_
So Bob is \_

The houses on Bank Street have odd numbers on one side and even numbers on the other side. Tom Smith lives at number 27.
What are the numbers on each side of Tom's house? \_\_and \_\_\

Ian is 6 feet tall.
David is one foot smaller than Ian. Bruce is two feet smaller than David.
So David is \_\_feet tall.
So Bruce is \_\_feet tall.

Add 10 to 35 \_\_
Doris is five years older than Grace. Grace is three years older than Edith.

Who is older, Doris or Edith? __________

How much older is she? __________ years.

Multiply 5 and 8 ________

Sandy has three jugs, one yellow, one green, and one white.

The yellow jug holds five pints less than the white one. The white jug holds four pints less than the green one.

Which holds more, the yellow jug or the green jug? __________

How much more does it hold? __________ pints.

Write nineteen in figures ________

Joy has seven pennies more than Olive. Olive has eight pennies more than Bess.

Who has more, Joy or Bess? __________

How many more has she? __________

A dozen eggs cost 4 shillings. How much do 2 dozen cost?

_______ shillings.
Archie jumps two feet less than Dennis. Dennis jumps three feet less than Peter.

Who jumps further, Archie or Peter? ____________

How much further does he jump? __________ feet.

---

Take 7 away from 19 __________

---

June has seven shillings more than Carol. Carol has five shillings more than Audrey.

Who has more, June or Audrey? ____________

How much more has she? __________ shillings.

---

Divide 9 by 3 __________

---

Joe has three books, one yellow, one black, and one orange.

The yellow book has ten pages less than the orange book. The orange book has five pages less than the black book.

Which has more pages, the yellow book or the black book? __________

How many more pages has it? __________ pages.

---

How many feet do four dogs have? __________ feet.
Cathy is five years older than Agnes. Agnes is eight years older than Jess.

Who is older, Cathy or Jess? __________

How much older is she? ________ years.

Add 4 to 9 _________

Bert has three marbles less than Charles. Charles has five marbles less than Jerry.

Who has more, Bert or Jerry? __________

How many more has he? ________ marbles.

If a dozen eggs cost 4 shillings, how much does one egg cost?

_________ pennies.
Problems used in investigation into asymmetry error:

1. **asymmetry forms**

A  Jean is 6. Jean is three years older than Sally.
What age is Sally? _______ years old.

Sp.  Write sixteen in figures ______

B  Eve has 6 pennies. Eve has one penny less than Sue.
How many pennies has Sue? _______ pennies.

Sp.  17-9 ______

C  Rose got up at 7 o'clock. Rose got up one hour after Betty
When did Betty get up? _______ o'clock.

Sp.  The houses on Bank Street have odd numbers on one side and even
numbers on the other side. Tom Smith lives at number 27.
What are the numbers on each side of Tom's house? _______ and_______

D  Jack weighs 7 stones. Jack weighs two stones less than Frank.
What weight is Frank? _______ stones.
E

May is 5 feet tall. May is one foot taller than Jill.

What height is Jill? ______ feet tall.

F

A can holds 4 pints of milk. The can holds one pint less than a jug.

How much milk will the jug hold? ______ pints.

G

Donald has 6 marbles. Donald has two marbles more than Ian.

How many marbles has Ian? ______ marbles.

H

Jean jumps 4 feet. Jean jumps two feet less than Ann.

What distance does Ann jump? ______ feet.

Sp.

In June one hen laid 28 eggs and another laid 26 eggs. How many eggs did the two hens lay between them? ______ eggs.

Sp.

One toy cost 6 shillings and another cost 7 shillings.

How many shillings did the two cost altogether? ______ shillings.
Colin has 4 pennies. Colin has two pennies more than Bill. How many pennies has Bill? ______ pennies.

The length of Tim's foot is 8 inches. Tim's foot is three inches shorter than Ian's. What is the length of Ian's foot? ______ inches.

Divide 15 by 5 ______

Tom weighs 8 stones. Tom is two stones heavier than Dick. What weight is Dick? ______ stones.


A can holds 5 pints of milk. The can holds two pints more than a jug. How much milk will the jug hold? ______ pints.
Multiply 6 and 9

Jill is 5. Jill is two years younger than Susan. What age is Susan? ________ years old.

Write fourteen in figures ________

The length of Jane's foot is 7 inches. Jane's foot is two inches longer than Pat's. What is the length of Pat's foot? ________ inches.

Divide 12 by 3 ________

Peter got up at 8 o'clock. Peter got up two hours before John. When did John get up? ________ o'clock.

Subtract 4 from 11 ________

Alice has 2 books, one red and one green. The red book has 60 pages. The red book has 10 pages more than the green book. How many pages has the green book? ________ pages.
Ian is 5 feet tall. Ian is one foot smaller than David. What height is David? _______feet tall.

Sp. Add 13 and 6 _______

Peter jumps 5 feet. Peter jumps two feet further than Andrew. What distance does Andrew jump? _______feet.


Jim has 7 marbles. Jim has 3 marbles less than Paul. How many marbles has Paul? _______marbles.

Sp. Add 10 to 35 _______
2. Non-asymmetry forms

A  Jean is 6. Sally is three years older than Jean.
What age is Sally? _______years old.

Sp. Write sixteen in figures _______ 

B  Eve has 6 pennies. Sue has one penny less than Eve.
How many pennies has Sue? _______pennies.

Sp. 17-9 _______ 

C  Rose got up at 7 o'clock. Betty got up one hour after Rose.
When did Betty get up? _______0'clock.

Sp. The houses on Bank Street have odd numbers on one side and even numbers on the other side. Tom Smith lives at number 27.
What are the numbers on each side of Tom's house? _____ and _____

D  Jack weighs 7 stones. Frank weighs two stones less than Jack.
What weight is Frank? _______stones.

Sp. 23-5 _______
E  May is 5 feet tall. Jill is one foot taller than May. What height is Jill? _____ feet tall.

F  A can holds 4 pints of milk. A jug holds one pint less than the can. How much milk will the jug hold? _____ pints.

G  Donald has 6 marbles. Ian has two marbles more than Donald. How many marbles has Ian? _____ marbles.


J  Colin has 4 pennies. Bill has two pennies more than Colin. How many pennies has Bill? _____ pennies.
K

The length of Tim's foot is 8 inches. Ian's foot is three inches shorter than Tim's.

What is the length of Ian's foot? _______ inches.

Sp.

Divide 15 by 5 _______

L

Tom weighs 8 stones. Dick is two stones heavier than Tom.

What weight is Dick? _______ stones.

Sp.

15-11 _______

M

Steven has 2 books, one blue and one yellow. The blue book has 50 pages.

The yellow book has ten pages less than the blue book.

How many pages has the yellow book? _______ pages.

Sp.

17-8 _______

N

A can holds 5 pints of milk. A jug holds two pints more than the can.

How much milk will the jug hold? _______ pints.

Sp.

Multiply 6 and 9 _______
Jill is 5. Susan is two years younger than Jill.

What age is Susan? _______ years old.

Write fourteen in figures _______

The length of Jane's foot is 7 inches. Pat's foot is two inches longer than Jane's.

What is the length of Pat's foot? _______ inches.

Divide 12 by 3 _______

Peter got up at 8 o'clock. John got up two hours before Peter.

When did John get up? _______ o'clock.

Subtract 4 from 11 _______

Alice has two books, one red and one green. The red book has 60 pages.

The green book has ten pages more than the red book.

How many pages has the green book? _______ pages.

25-24 _______

Ian is 5 feet tall. David is one foot smaller than Ian.

What height is David? _______ feet tall.
Add 13 and 6

Peter jumps 5 feet. Andrew jumps two feet further than Peter. What distance does Andrew jump? _______ feet.

One book costs 1\text{\$} shillings and another costs 15 shillings. How many shillings did the two cost altogether? _______ shillings.

Jim has 7 marbles. Paul has three marbles less than Jim. How many marbles has Paul? _______ marbles.

Add 10 to 35 _______
Problems combining overlap and asymmetry difficulties

OA/A**
Jean is 7.
Jean is three years older than Sally. Jean is one year older than Linda.

So Sally is ___________ years old.
So Linda is ___________ years old.

Sp. Write sixteen in figures ___________ 

OA/B
A jug holds 5 pints of water.
The jug holds two pints less than a bucket. The jug holds one pint less than a can.

So the bucket holds ___________ pints.
So the can holds ___________ pints.

Sp. Multiply 7 and 8 ___________

OA/C
Donald has 9 marbles.
Donald has three marbles more than Ian. Donald has four marbles more than Brian.

So Ian has ___________ marbles.
So Brian has ___________ marbles.

* See pages
** Where OA/A means overlap form, asymmetry version/Question A
One toy cost 6 shillings and another cost 7 shillings. How many shillings did the two cost altogether? ________ shillings.

Joyce jumps 5 feet. Joyce jumps one foot less than Ann. Joyce jumps three feet less than Nancy.

So Ann jumps ________ feet.

So Nancy jumps ________ feet.


Colin has 12 shillings. Colin has five shillings more than Bill. Colin has three shillings more than Eric.

So Bill has ________ shillings.

So Eric has ________ shillings.

Take 7 away from 19 ________
Alice has three books, one red, one green and one blue. The red book has 60 pages. The red book has twenty pages less than the green book. The red book has ten pages less than the blue book.

So the green book has ______ pages.
So the blue book has ______ pages.

---

How many feet altogether do four dogs have? ______ feet.

---

Alan is 7.
Alan is one year older than Tony. Alan is four years older than Bob.

So Tony is _______ years old.
So Bob is _______ years old.

---

If 12 bars of chocolate cost 6 shillings, how much does one bar cost?

_______ pennies.

---

In Ian's house, the kitchen is 10 feet long.
The kitchen is one foot shorter than the bedroom. The kitchen is two feet shorter than the living-room.

So the bedroom is _________ feet long.
So the living-room is _________ feet long.
Add 4 and 9

Jean is 7.
Jean is three years older than Sally. Sally is one year older than Linda.
So Sally is years old.
So Linda is years old.

Write sixteen in figures

A jug holds 5 pints of water.
The jug holds two pints less than a bucket. The bucket holds one pint less than a can.
So the bucket holds pints.
So the can holds pints.

Multiply 7 and 8

Donald has 9 marbles.
Donald has three marbles more than Ian. Ian has four marbles more than Brian.
So Ian has marbles.
So Brian has marbles.

* Where NOA/A means non-overlap form, asymmetry version/Question A
Sp.

One toy cost 6 shillings and another cost 7 shillings.
How many shillings did the two cost altogether? ______ shillings.

NOA/D

Joyce jumps 5 feet.
Joyce jumps one foot less than Ann. Ann jumps three feet less than Nancy.

So Ann jumps ___________ feet.
So Nancy jumps ___________ feet.

Sp.

A dozen eggs cost 4 shillings. How much do 2 dozen cost?
__________ shillings.

NOA/E

Colin has 12 shillings.
Colin has five shillings more than Bill. Bill has three shillings more than Eric.

So Bill has ___________ shillings.
So Eric has ___________ shillings.

Sp.

Take 7 away from 19 ____________

NOA/F

Alice has three books, one red, one green and one blue.
The red book has 60 pages.
The red book has twenty pages less than the green book. The green book has ten pages less than the blue book.

So the green book has ___________ pages.
So the blue book has ___________ pages.
How many feet altogether do four dogs have? ________ feet.

Alan is 7.
Alan is one year older than Tony. Tony is four years older than Bob.

So Tony is ________ years old.
So Bob is ________ years old.

If 12 bars of chocolate cost 6 shillings, how much does one bar cost? ______________ pennies.

In Ian's house, the kitchen is 10 feet long.
The kitchen is one foot shorter than the bedroom. The bedroom is two feet shorter than the living-room.

So the bedroom is ________ feet long.
So the living-room is ________ feet long.

Add 4 and 9 ________
Problems for study in Appendix I. pages 152 ff.
(Note: non-overlap forms as for original overlap error investigation)

0/J
Grace is five years older than Doris. Grace is three years older than Edith.
Who is older, Doris or Edith? __________
How much older is she? ___________ years.

0/Sp.
Multiply 5 and 8 __________

0/K
Sandy has three jugs, one yellow, one green, and one white.
The white jug holds five pints less than the yellow one. The white jug holds four pints less than the green one.
Which holds more, the yellow jug or the green jug? __________
How much more does it hold? __________ pints.

0/Sp.
Write nineteen in figures __________

0/L
Olive has seven pennies more than Joy. Olive has eight pennies more than Bess.
Who has more, Joy or Bess? __________
How many more has she? __________ pennies.

Dennis jumps two feet less than Archie. Dennis jumps three feet less than Peter.
Who jumps further, Archie or Peter? __________
How much further does he jump? __________ feet.

Take 7 away from 19 __________

Carol has seven shillings more than June. Carol has five shillings more than Audrey.
Who has more, June or Audrey? __________
How much more has she? __________ shillings.

Divide 9 by 3 __________

Joe has three books, one yellow, one black and one orange.
The orange book has ten pages less than the yellow book. The orange book has five pages less than the black book.
Which has more pages, the yellow book or the black book? __________
How many more pages has it? __________ pages.

How many feet do four dogs have? __________ feet.
C/P
Agnes is five years older than Cathy. Agnes is eight years older than Jess.

Who is older, Cathy or Jess? __________
How much older is she? __________ years.

C/Sp.
Add 4 to 9 __________

C/Q
Charles has three marbles less than Bert. Charles has five marbles less than Jerry.

Who has more, Bert or Jerry? __________
How many more has he? __________ marbles.

C/Sp.
If a dozen eggs cost 4 shillings, how much does one egg cost?
______________ pennies.
REFERENCES

Atkinson, R.C. (1963): 'Mathematical models in research with children';


Boring, E.G. (1950): 'A history of experimental psychology'.

Bruner, J. (1957): 'Going beyond the information given'; in Gruber, H.E.
   et al., Pp. 41-69 (q.v.).

Bruner, J. (1962): 'On knowing: essays for the left hand'. Cambridge:
   Harvard University Press.

   Clark.


Chapman, L.J. and Chapman, J.P. (1959): 'Atmosphere effect re-examined'.
   J. Exp. Psychol., 58, 220-226.


Colby, K.M. (1963): 'Computer simulation of a neurotic process'; in

Denny-Brown, D. (1932): 'Theoretical deductions from the physiology of the cerebral cortex'; quoted in Hebb, D.O., 1949, P.7 (q.v.).

Donaldson, M.C. (1956): 'The relevance to the theory of intelligence testing of the study of errors in thinking'.

Thesis submitted for the degree of Ph.D. in the University of Edinburgh.


Festinger, L. (1957(a)): 'The relation between behavior and cognition'; in Gruber, H.E. et al. (q.v.).

Festinger, L. (1957(b)): 'A theory of cognitive dissonance'.


Flavell, J.H. (1963): 'The developmental psychology of Jean Piaget'.


Goldstein, K. and Scheerer (1941): 'Abstract and concrete behavior: an experimental study with special tests'. Psychol. Monogr., 53, 151PP.

Gruber, H.E. et al. (the cognition symposium committee) (1957): 'Contemporary approaches to cognition' ('The Colorado symposium'). Cambridge: Harvard Univ. Press.


Harlow, H.F. (1958): 'Behavioral contributions to interdisciplinary research'; in Harlow, H.F. and Woolsey, C.N. (eds.).


Hebb, D.O. (1958): 'Alice in Wonderland or psychology among the biological sciences'; in Harlow, H.F. and Woolsey, C.N. (eds.) (q.v.).


Hebb, D.O. (1961): Introduction to article by W. Heron; in Solomon, F. et al. (eds.) (q.v.).
Holt, R.R. (1950): 'Some statistical problems in clinical research'.
Educ. and psychol. measurement, 10, 609-627; reprinted in Beck, S.J.
and Molish, H.B., 1959, Pp. 512-519 (q.v.).


Hunt, J. McV. (1963); in Harvey, O.J.: 'Motivation and social inter-
action': a symposium. N.Y. Ronald Press.

Hunter, I.M.L. (1957(a)): 'The solving of three-term series problems'.

Hunter, I.M.L. (1957(b)): 'Note on an atmosphere effect in adult reasoning'.

Hunter, W.S. (1912): 'The delayed reaction in animals and children'.
Behav. Monogr., 2 (1), 1-85.

Hunter, W.S. (1918): 'The temporal maze and kinaesthetic sensory processes
in the white rat'. Psychobiol., 2, 339-351.

Inhelder, B. (1962): 'Some aspects of Piaget's genetic approach to cognition'
in Kessen, W. and Kuhlman, C., Pp. 19-34 (q.v.).

Inhelder, B. and Piaget, J.: 'The growth of logical thinking from child-


Koch, S. (1959(a)): 'Psychology, a study of a science, Vol. 1: sensory, perceptual and physiological formulations'.

Koch, S. (1959(b)): 'Psychology, a study of a science, Vol. 2: General systematic formulations, learning and special processes.'


Piaget, J. (1921): 'Une forme verbale de la comparaison chez l'enfant'. Arch. Psychol., 18, 141-172.


Piaget, J. (1953(a)): 'Logic and Psychology'. Manchester: Manchester University Press.


Piaget, J. (1963): Foreword to Flavell (q.v.).


Pribram, K.H. (1960): 'A review of theory in physiological psychology'.  

Reitman, W.R. (1963): 'Personality as a problem solving coalition';  
in Tomkins, S.S. and Messick, S., Pp. 69-99 (q.v.).

Scott, N.C., Jr. (1964): 'Science concept achievement and cognitive  
functions'. J. Res. in Sci. Teaching, 2 (1), 7-16.


Simon, B. (1957): 'Psychology in the Soviet Union'. London:  
Routledge and Kegan Paul.

Simon, B. and Simon, J. (eds.) (1963): 'Educational psychology in the  

development'; in Kessen and Kuhlman, Pp. 150-155 (q.v.).

Slavina, L.S. (1954): 'Specific features of the intellectual work of  
unsuccessful pupils'; in Simon, B., 1957, Pp. 205-212 (q.v.).

Smirnov, A.A. (1955): 'Psychological research, 1953-1955'; in Simon,  
1957, Pp. 29-45 (q.v.).

'The lack of cerebral effects of d-tubocurarine'. Anesthesiology, 8, 1-14.


Tomkins, S.S. (1963): 'Simulation of personality: the interrelationships between affect, memory, thinking, perception and action'; in Tomkins and Messick, Pp. 3-57 (q.v.).


