REDUNDANT AND IRRELEVANT DATA IN PROBLEM SOLVING

by

ALISON CHALMERS CAMPBELL

Thesis presented for the Degree of Doctor of Philosophy

of the University of Edinburgh

in the Faculty of Social Sciences

November, 1964

TABLE OF CONTENTS

REDUNDANT AND IRRELEVANT DATA IN PROBLEM SOLVING

I. INTRODUCTION

A general feature of psychological studies of human problem solving is that the subject is given material sufficient for solution of the problem, and no more than this. Thus in one of the problem situations devised by Maier (1930) the subject is presented with two clamps, a piece of chalk, a length of wire, and three lengths of wood. There is also a heavy table that is not to be moved, plus the four walls and low ceiling of a room. The problem is to construct two pendulums to swing over and mark two specified spots on the floor. The given materials are adequate for solution of the problem: nothing more than them is required. At the same time all of them are required. Likewise with Durkin's (1937) puzzles requiring a Greek cross to be assembled from a number of irregularly shaped component pieces: no pieces are given other than those actually required. Neither in these nor other similar studies does the question arise of whether one could - or should - dispense with part(s) of the data. Attention has been focussed on the influence of experiential and motivational factors on problem solving activities, and subjects have been nurtured in the tradition that all the data are necessary and are there to be used.

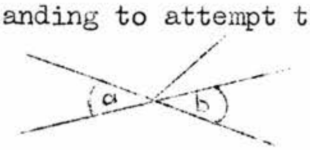
With test items also, which can be regarded as problem situations in miniature, it is standard practice to present the subject with data sufficient for arriving at solution and no more (nor any less) than these. Consider opposites, mixed sentences, definitions, synonyms, analogies, and similar items from tests of intellectual ability. No superfluous data are given; no demand is made for selection of data relevant to item solution plus rejection of those not relevant. Selection may be called for in deciding between multiple choice alternatives, but this is another matter entirely. Similarly for more complex types of item, e.g.

> I have five friends who go to the same school as I do. Let us call them A, B, C, D, and E.
> A and C are boys; the rest are girls.
> C and E have fair hair; the others have dark hair.
> B and C are in the class below me; the others are in my class.
> Now answer the following questions by underlining in the brackets the correct answer to each.
> (1) Who is the boy in my class ? (A/B/C/D/E)
> (2) Who is the fair-haired girl ? (A/B/C/D/E)
> (3) Who is the boy in the class below me ? (A/B/C/D/E)
> (4) Who is the dark-haired girl in my class ? (A/B/C/D/E)
> (5) Who is the dark-haired girl in the class below me ? (A/B/C/D/E)

This item comes from one of the Moray House Verbal Reasoning Tests. To answer questions (1)- 5) correctly one must "perceive the inter-relations of the three separate dichotomies" and take "the several cross-classifications" into account (Pilliner, 1961, p.67). All the data are brought into play at some stage, and none can be dispensed with.

The tradition that all the data provided are there to be used is strongly reinforced by the educational practice of presenting the student with orderly and organised arrays of facts and principles with all irrelevant and extraneous matter excluded. Polya (1957) repeatedly exhorts the person who has struck difficulty with a mathematical problem to consider whether all the data have been taken into account; and formal logic teaches one to work from what is given, all of it, and it alone - as with the syllogism. Luchins and Luchins (1950) mention high school students who had learned in their geometry class to check off each piece of information as used, and if any were left unchecked at the end to regard this as a sign of not having proceeded correctly. Should it not be possible for all irrelevant and extraneous matter to be excluded from educational material, devices such as diagrams and models are used to make what is relevant stand out from an (otherwise) embedding background. Thorndike (1931) advocated that relevant features be made as prominent as possible, and later writers argue likewise. Kossov (1960) notes that whenever possible the essentials of a situation are highlighted by removing extraneous detail. His own concern is with techniques for highlighting the crucial aspects of educational material when actual removal of extraneous detail is not possible.

Small wonder, in view of educational practice over the years, that subjects tend to assume that something is wrong if they have not made use of all the data provided. But to foster such an assumption has its dangers, for it can be a sign of weakness of intellect and of defective understanding to attempt to use all the data. If one is given the figure



and asked to prove that $a = b$, the sensible thing is to ignore the fifth line completely (Wertheimer, 1945). The hallmark of "sensible" (as opposed to "senseless") procedures is that they do not take all items into account, but all relevant items. The concentration on and presenting of only such materials as are pertinent to the issue at hand has been attacked by Wertheimer (1945), Luchins & Luchins (1950), Bruner (1960) and Abercrombie (1962), among others. Luchins and Luchins suggest that problems involving the discovery, selection and evaluation of facts and hypotheses should be introduced into all school studies, and Bruner advocates school curricula that will allow the pupil actively to discover, select and evaluate data. Active coping with the data is also emphasised in the schemes for fostering "creativity", currently popular in the United States (e.g. Taylor, 1964). These schemes are regarded as remedial, and necessary because school studies are presented in too cut-and-dried a manner and so fail to develop a sensitivity to problems and problem structure. Despite these attacks, suggestions, and remedial schemes, the recently developed techniques of programmed learning continue to perpetuate the tradition that all the data provided are there to be used. Programmes focus unerringly on the "crucial features" of the learning material, and that this is desirable seems not seriously to have been questioned until Pressey (1963), a pioneer in the field of automated instruction, pointed out that human learning is largely an integrative and judgmental activity involving a search for and selection of relevant aspects of the data, and that programming as currently carried out overlooks this completely and is therefore unsatisfactory. Wohlwill (1962) and Smedslund (1964) also express discontent on this score, and Thouless (1963) considers that teaching programmes are reduced to instruments of indoctrination rather than education unless they require the subject to evaluate the data presented. Calder (1964) goes so far as to suggest that data irrelevant to the questions asked might usefully be incorporated into a learning programme in order to stimulate the learner's critical faculties and keep him at the task. These are, however, the voices of the avant-garde. In general it is accepted without question that the subject is presented with data sufficient for the task, and these alone.

Problems as met with in the "real life" situation are never as clearly defined as are those presented to the pupil in the schoolroom, or to the subject in the psychological laboratory. We are at all times enmeshed in an enormous welter of data, and stimuli impinge upon us from all sides. Our efficiency in living our daily lives depends on how we deal with this bombardment of information. There being no benevolent teacher or experimenter standing by to present us with just the materials required for solving any given problem, one needs to be skilled in "discarding all irrelevancies, in going for the jugular of the problem" (Posner, 1962, p.10)*. . Polya (1957, p.152) is careful to point out that the question to be asked in the real life situation, in contrast to that of solving mathematical problems, is not "Have I used all the data ?" but rather "Have I used all the data that could contribute appreciably to the solution ?" The business of selecting what are relevant from among the countless data that confront us, and disregarding all the rest, is a prime obstacle to problem solving in everyday life. It is a prime obstacle also to scientific advance. Various features attendant on different observations of a phenomenon will be more or less irrelevant to its occurrence, but the discovery of these features is not as simple as one might at first think. It took a Galileo to recognise that some of the most obvious differences between a feather and a stone are irrelevant to how quickly each will fall from a given height.

One cannot in the laboratory deal with masses of material as extensive as those with which we are confronted in everyday life, and it has been implied by some (e.g. Woodworth, 1938, p.772; Hebb, 1958, p.211)

---

* Quoted from GERARD, R. W., Neuro-physiology: an integration. In <u>Handbook of Physiology. Volume III</u>, American Physiology Society, 1960.

that if the problem solving process is to be brought within the confines of the laboratory then the supplying of the subject with only such materials as are needed to solve the problem is a device that has to be used. This can however be challenged. It is possible to inject a problem situation with limited amounts of data over and above those logically sufficient for attaining the solution, while still having the situation under experimental control. It is this that the present research seeks to do. The additional data will, as regards their logical status, be either relevant or not relevant to solution. The distinction between these two sorts of additional data is taken up in Section II. Suffice it here to say that additional relevant data are data that are able to be used in arriving at solution, but their utilisation is not necessary. They are superfluous and introduce redundancy. Additional irrelevant data are such as have no bearing on solution, and cannot be used in arriving thereat.* By noting how these different sorts of additional data are coped with, and how they influence the course of problem solving, at least a partial understanding may be achieved of what goes on when, in the real life situation, problems are solved in the face of great masses of data that are largely superfluous or misleading.

---

* It is not to be assumed that redundant and irrelevant data are always "additional" to what are logically sufficient for solution: both can be present without there being sufficient data for a determinate solution to be reached. Such situations are not however the concern of the present research.

II. PROBLEMS USED IN THE RESEARCH

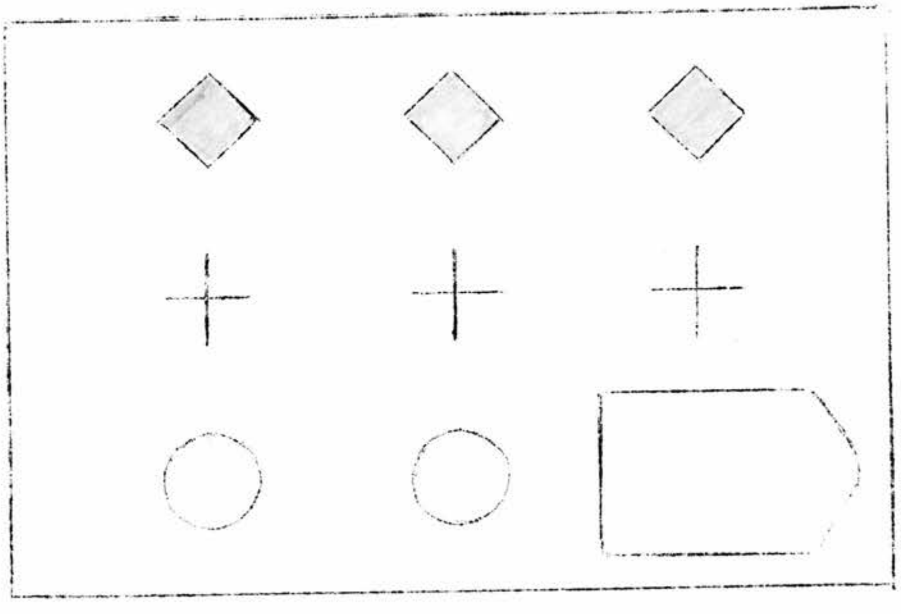
Test items were chosen as a type of problem situation eminently suited to empirical study. The structural elements that go to make up an item can be precisely specified and controlled, and no complicated apparatus is required. The study of test items is also to be encouraged in that it may lead to improvements in test construction. To assess the effect on problem solving of data in excess of those logically sufficient for arriving at the designated solution, test items are needed into which additional relevant data and additional irrelevant data can both be introduced. To assess the possible interaction of these two sorts of additional data is also desirable, since both are simultaneously present in everyday situations. Hence the items must be such that the two sorts of additional data can be introduced not only separately but also in conjunction. No suitable item type was immediately available, and to discover one was no simple matter. While a few items in published tests of intellectual ability do depart from standard practice and present the subject with data over and above those logically sufficient for arriving at the designated solution, the extra data are either relevant or irrelevant but never both. In these items from the Shipley-Hartford test, for example, the solution is overdetermined and the data are therefore redundant:

mist is	wasp as	pint in	tone - -
12321	23432	34543	456 - -

It is difficult however to see how words or numbers irrelevant to solution could be incorporated into these items without destroying their basic structure. Certain items from Sets C, D and E of Raven's

Progressive Matrices (1938) likewise have data more than sufficient to determine the answer, e.g.

D1



The first and second columns duplicate each other and one or other could be omitted without rendering the answer unattainable. Alternatively, either of the first two rows could be omitted since one alone suffices to illustrate the constructional principle on which the matrix is based. Indeed, for as simple a matrix as this, both a row and a column could be omitted without the answer becoming ambiguous. But while the matrix is, as it stands, redundant, it is again difficult to see how data (figures) irrelevant to solution could be brought into the item without destroying its basic structure. Items containing irrelevant data are rare, writers on test construction (e.g. Adkins, 1947; Ebel, 1951; Bean, 1953; Downie, 1958) having laid it down almost as an unbreakable rule that all extraneous and unrelated detail should be eliminated. Bean (p.67) does however allow that there may be some call for items in which the subject has to pick out relevant from irrelevant facts, and such items are found in tests of arithmetical reasoning, e.g.

> 6 typists are hired at 5/- an hour to type out 5672 application forms. They work at an average rate of 36 forms per hour, and they complete the whole task in 4 working days. At the same speed of work, how many typists would be needed to complete the same task in 2 working days ?

What is underlined is mere "padding". Items with irrelevant data, while appearing but infrequently in published tests, have been used from time to time for research purposes, e.g.

> A. Boats A and B each have 5 gallons of fuel.
> B. They cruise together towards the same point.
> C. Boat A uses 1 gallon of gas going 3 miles.
> D. Boat B uses 2 gallons of gas going 1 mile.
> E. Boat B is twice as heavy as boat A.
>
> Question: How many miles apart will the two boats be when they run out of gas ?

This is an item type devised by Guilford and his colleagues (1950). One of the statements (here E) is irrelevant to the answering of the question. The irrelevant statement is to be indicated, but the question itself is not worked. Redundant data could conceivably be put into the above two item types, in conjunction with the irrelevant. The items do not however allow as precise a control over the different structural elements as was considered desirable for the purposes of the present research.

Any research, in its actual carrying out, tends to be a "rather informal, often illogical and sometimes messy-looking affair" with a "great deal of floundering around in the empirical world, sometimes dignified by names like 'pilot studies' or 'exploratory research'" (Taylor et al., 1959, p. 169). Considerable floundering around preceded the discovery of an appropriate item type, and pilot studies were run using nonverbal matrices, number series, and letter series items, none of which proved wholly satisfactory. Finally a type of code item was found which can be so manipulated as to contain additional relevant data; additional irrelevant data; both; or neither; and with the stimulus elements precisely specified and controlled. The items consist of five 5-letter words listed across the page, with coded versions of the words listed vertically and in a different (randomised) order underneath, e.g.

(1)CLOUD (2)COUNT (3)CIDER (4)CHAIN (5)CARTS

```
- - J - P    ..............( 4 )
- - - P -    ..............( 2 )
- J Z - -    ..............( 5 )
- - - - Z    ..............( 3 )
- - - - -    ..............( 1 )
```

Item I

The code is arbitrary, in the sense of not being based on any rule of alphabetical sequence, and the problem is to locate the coded counterparts of the words (1)-(5) and insert the appropriate numbers in the pairs of brackets. The solution has been indicated.

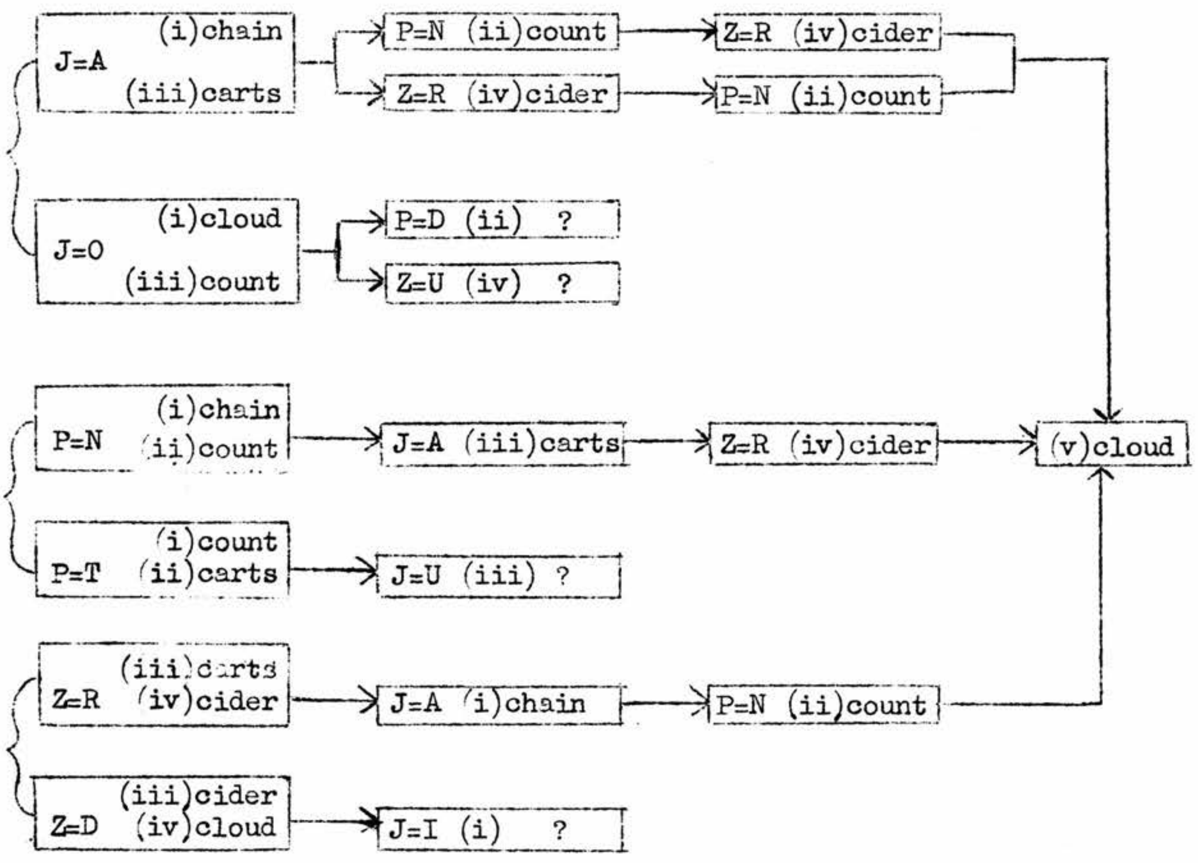
Although the coded words are incomplete, the given pairs of coded letters are such as to determine a unique and unequivocal solution. To discover which incomplete coded word corresponds to each of the words (1)-(5) one has to consider the positions in which the different letter pairs occur. J's for example occur in third and second positions in the first and third coded words respectively. Hence these two coded words represent a pair of words of which the first has a certain letter in third position while the other has this same letter second. Two possibilities arise:

(i) CLOUD and COUNT, with J standing for O

(ii) CHAIN and CARTS, with J standing for A.

Suppose - - J - P : represents CLOUD. P would then stand for D, and the second coded word would represent a word with D in fourth position. No such word is listed. Hence - - J - P cannot be CLOUD and must instead be CHAIN (with the third coded word as CARTS). If - - J - P is CHAIN then P = N. Hence the second coded word represents a word with N in fourth position. Only COUNT fits this requirement. The third coded word being CARTS, Z = R. The fourth coded word must therefore represent a word ending in R, viz. CIDER. The coded word for which no letters are given must, by default, be CLOUD. It is the only word left.

For Item I the given letters are, in toto, just sufficient for the solution to be attained. All have to be used, although not necessarily in the sequence outlined above. The various routes to solution are shown in the accompanying flow chart:



Flow Chart for Item I

The numbers (i)-(v) in the flow chart overleaf refer to the coded words as listed in order down the page. The item may be regarded as a maze for which there are three starting points that open up correct routes (really variants on a single theme) plus three starting points that are false leads and result in dead ends. In drawing this and other flow charts it has been assumed that information acquired at one step will be utilised at the subsequent step. Thus it is here assumed that if the subject begins by establishing that P = N then his next step will be to look at the J's rather than the Z's, since a P and a J occur together in the first coded word whereas there is no overlap of P's and Z's. Most subjects do in fact work in this way, and while the little-used branch-routes to solution (e.g. from P's to Z's to J's) could be shown in the flow charts, to do so would make the charts excessively and unnecessarily complex. In Item I there are two initially-plausible identifications for each coded letter pair. The items used in the research ranged from those with a single possible identification for each letter pair (i.e. no false leads) to those with four initially-plausible identifications for each letter pair and hence a probability of 3/4 = 0.75 of starting out on a wrong track. Flow charts for this latter sort of item become quite complex.

It might be noted that although all the coded letters in Item I must be used if the solution is to be validly attained, none of these letters is "necessary" in an absolute sense. Alternative subsets of letters can be found that are equally as appropriate, e.g.*

---

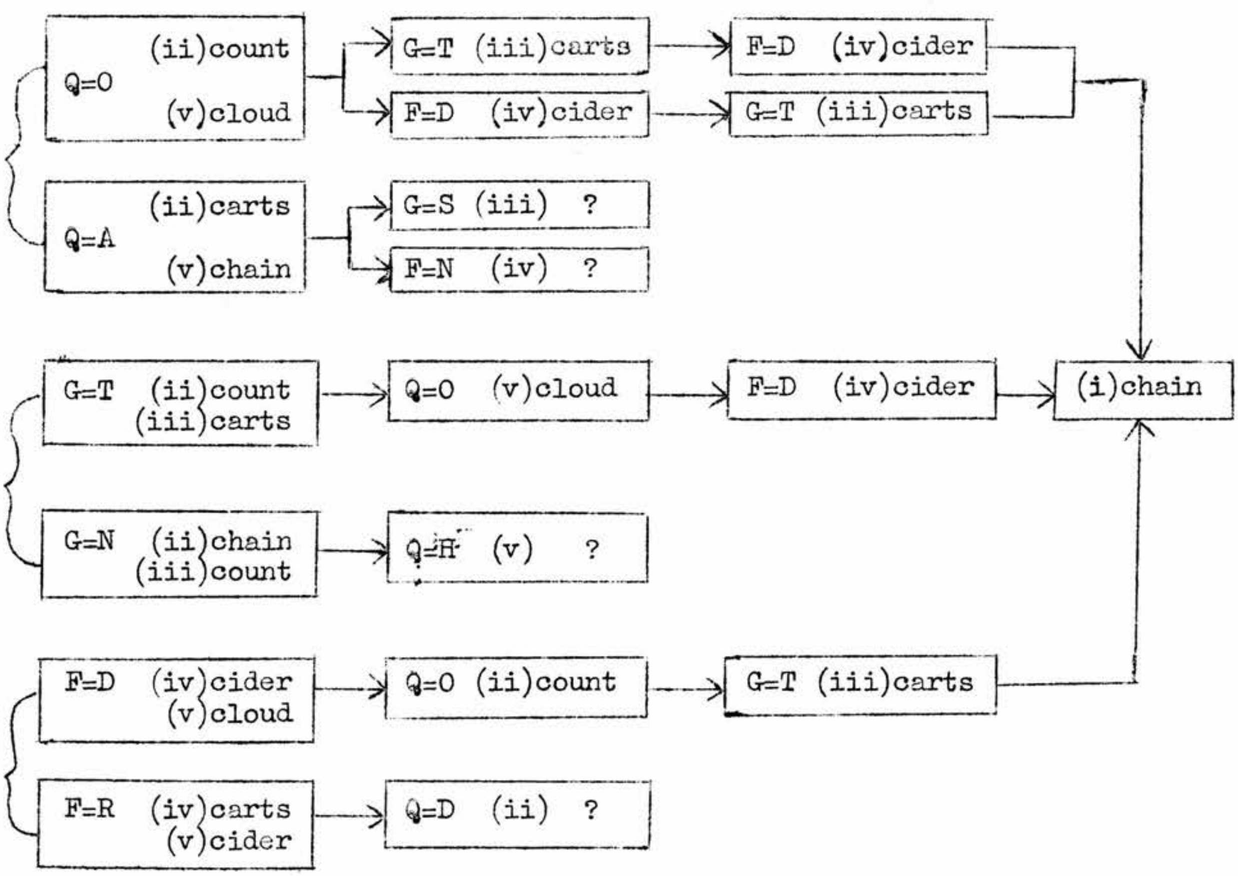
* Codes are by no means the only item type where the data provided, while needed as they stand, are not necessary in an absolute sense. Take for example the verbal analogy item Up : Down = In : ? Any word and its opposite could be substituted for Up : Down, and would be equally as appropriate.

```
(1)CLOUD   (2)COUNT   (3)CIDER   (4)CHAIN   (5)CARTS

        -  -  -  -  -     .............(   )
        -  Q  -  -  G     .............(   )
        -  -  -  G  -     .............(   )
        -  -  F  -  -     .............(   )
        -  -  Q  -  F     .............(   )
```

Item I'

The different routes to solution for Item I' fall into the same sort of pattern as for Item I:



Flow Chart for Item I'

It might also be noted that the just-sufficient data in this sort of code item are not necessarily minimally sufficient. Different numbers of letter pairs can be "just sufficient" for arriving at solution, depending on which particular pairs they are. As a simplified example take the following:

```
      (1) AREA          (2) ACRE          (3) RACE

          -  -  -  -       ............(   )
          J  -  -  -       ............(   )
          -  -  J  -       ............(   )


          -  -  -  Q       ............(   )
          -  M  -  -       ............(   )
          M  -  -  Q       ............(   )
```

In the first arrangement the pair of J's fixes the solution. Only one pair of words - RACE and ACRE - has a first and third letter the same. Hence the second coded word is RACE and the third ACRE, which leaves the first to be AREA. In the second arrangement alternative possibilities arise, and neither the M's nor the Q's alone suffices to determine an unambiguous solution. In the interests of standardisation, code items with just sufficient data were always so structured as to have three letter pairs, all needed if solution is to be unambiguous. The letter pairs are always arranged so as to have two coded words with two coded letters, two coded words with one coded letter, and one coded word without any coded letters (as in Items I and I' above). Never do the two coded words with two coded letters both have the same two coded letters, i.e. patterns such as

```
- Q - - G
- - G Q -
```

are avoided.

Items with just sufficient data provide a baseline for comparison with other versions which contain additional data.

Schematically the possibilities are as follows:

		Additional relevant (redundant) data	
		Absent	Present
Additional irrelevant data	Absent	1	2
	Present	3	4

Cell 1 represents items with just sufficient data for solution, all relevant and all to be used. Cell 2 items have additional relevant data; cell 3 items have additional irrelevant data; and cell 4 items have additional data of both sorts. The items with additional relevant data were so constructed as to have three letter pairs over and above those in the versions with just sufficient data, e.g.

(1)CLOUD (2)COUNT (3)CIDER (4)CHAIN (5)CARTS

```
- - J - P  ..............(   )
- Q - P G  ..............(   )
- J Z G -  ..............(   )
- - F - Z  ..............(   )
- - Q - F  ..............(   )
```

Item II

Again, never do two different coded letters (e.g. a J and a P) appear together in two different coded words. In Item II not all the given letters need be used. Sufficient for solution are the J's, P's and Z's (from Item I); the Q's, G's and F's (from Item I'); and various other subsets of the six letter pairs as well. The solution is considerably overdetermined, and parts of the data redundant. Items of this sort might be compared to Dashiell's (1930) checkerboard or open-alley mazes, which offer many alternative routes from the entrance to the goal. Some may wish to argue that if certain of the letters in Item II are not used by a subject in arriving at his solution then these letters are "irrelevant" rather than "relevant but redundant". The answer to this is that whether letters are relevant or not relevant to solution is here being judged in terms of the logical structure of the item - not in terms of which letters the subject actually uses (or tries to use). "Redundant" letters are always potentially useful; "irrelevant" letters never so. If the original letter pairs from Item I (J's, P's and Z's) are regarded as a baseline, Item II is 50% redundant, i.e. 50% of the letters are superfluous. It would be neat if one could take <u>any</u> set of three letter pairs and say "these are sufficient for solution and the rest are not needed", but this it was not possible to achieve. The J's, P's and G's are not, for example, sufficient for solution of Item II: unless another pair of letters is also taken into account, the identity of the fourth and fifth coded words remains in doubt. While an inauspicious choice of coded letter pairs may, as with the J's, P's and G's, necessitate the taking into account of more than three letter pairs, can an auspicious choice ever lead to fewer than three letter pairs being sufficient to determine a unique and unequivocal solution ? Only if there is but a single initially-plausible identification for each letter pair, in which case two pairs, provided they are non-overlapping, can suffice to fix solution. Two non-overlapping pairs of letters may <u>seem</u> sufficient also for items where there is more than one initially-plausible identification per letter pair, but this is not in fact so. Take the P's and F's of Item II. Equating P with N and F with D would establish the first, second, fourth and fifth coded words as CHAIN, COUNT, CIDER and CLOUD respectively. This leaves the third coded word to be CARTS. This solution is in fact "correct", and it has been arrived at from the P's and F's alone. It is not however the <u>only</u> solution to which the P's and F's, if taken alone, can lead. If P is equated with T instead of N - with F still equal to D - this establishes the first and second coded words as COUNT and CARTS, the fourth and fifth coded words remain as CIDER and CLOUD, and the third will now be CHAIN. For solution of Item II to be unambiguous, at least three letter pairs must be considered. The drawing of detailed flow charts for items with extra (redundant) letter pairs was not undertaken Whichever letter pair is chosen as starting point, there are multiple branch-routes stemming therefrom. Unlike the items with just sufficient data, there is no knowing which letter pair a subject will move on to next. To list all the possible branch-routes to solution would serve no useful purpose, and the possible starting points alone were listed. Thus for Item II there are six letter pairs, each with two initially-plausible identifications, which gives a total of twelve possible starting points. Six of these are false leads - those without an asterisk in the list below.

J=A	(i)chain (iii)carts	*	Q=0	(ii)count (v)cloud	*
J=0	(i)cloud (iii)count		Q=A	(ii)carts (v)chain	
P=N	(i)chain (ii)count	*	G=T	(ii)count (iii)carts	*
P=T	(i)count (ii)carts		G=N	(ii)chain (iii)count	
Z=R	(iii)carts (iv)cider	*	F=D	(iv)cider (v)cloud	*
Z=D	(iii)cider (iv)cloud		F=R	(iv)carts (v)cider	

Possible starting points for Item II

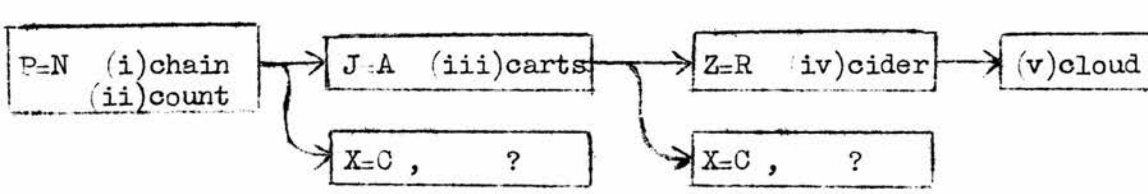
Items with additional irrelevant data were arrived at in two different ways. For items such as Item I the procedure is as follows:

(1)CLOUD	(2)COUNT	(3)CIDER	(4)CHAIN	(5)CARTS

X	-	J	-	P	....	( )
-	-	-	P	-	....	( )
X	J	Z	-	-	....	( )
X	-	-	-	Z	....	( )
X	-	-	-	-	....	( )

Item III

The X's are irrelevant to solution. Knowing that X = C is of no help in deciding which uncoded word is the counterpart of each coded word, since all the uncoded words begin with C. The ordinal position of the irrelevant letters was varied from item to item, i.e. in another item all the second letters might be the same, or all the third letters or fourth letters or fifth letters. A flow chart for Item III will be basically the same as for Item I, but with the addition at every step of dead ends stemming from the X's. Thus the route to solution starting from P = N could be represented as follows:



The first [X=C , ?] implies that, X being equal to C,

(iii) = cloud or cider or cards
and
(iv) = cloud or cider or cards
and
(v) = cloud or cider or cards

The second [X=C , ?] implies that, X being equal to C,

(iv) = cloud or cider
and
(v) = cloud or cider

Should one begin solution of the item with an X, the possibilities are thus:

X=C, (i)cloud, count, cider, chain or carts

X=C, (iii)cloud, count, cider, chain or carts

X=C, (iv)cloud, count, cider, chain or carts

X=C (v)cloud, count, cider, chain or carts

No real headway can be made here, and the subject soon seeks another starting point.

In contrast to Item III is Item IV:

(1)FOLKS (2)TABLE (3)DRIFT (4)LIKEN (5)STEAK

-	G	-	C	-	....	( )
-	R	-	-	-	....	( )
-	H	-	M	-	....	( )
C	V	-	-	U	....	( )
U	J	M	-	-	....	( )

Item IV

The letters in second position are irrelevant to solution: each of them is a "dead end", since knowing what any one of them stands for does not assist in the identification of any further word. Instead of the irrelevant letters all being the same they are in this case all different, both from each other and from the letters of the letter pairs. With this sort of item, too, the ordinal position of the irrelevant letters was varied from item to item. Item V illustrates yet a further development:

(1)VINYL (2)MASKS (3)CATER (4)PLUME (5)BRAWN

-	J	X	F	-	..............( )
-	X	-	-	J	..............( )
-	-	-	Z	-	..............( )
D	G	-	-	-	..............( )
H	-	-	-	G	..............( )

Item V

There are three letter pairs - X's, J's and G's - plus four single letters. The four single letters are irrelevant to solution, but they here occupy two different ordinal positions (first and fourth) instead of only one. For all the items containing irrelevant letters care was taken to have the letters deemed "irrelevant" not only without any positive bearing on solution but also without any negative (indirect) relevance thereto. The following would not for example be acceptable:

(1)GRAIN (2)PYLON (3)SKITE (4)ANGLE (5)GULPS

-	-	D	W	-	..............( )
D	-	Q	F	Z	..............( )
-	-	-	C	-	..............( )
-	-	-	X	Z	..............( )
Q	-	-	B	-	..............( )

None of the letters in fourth position has any positive bearing on solution. But once it has been established that Q - - B - represents GULPS, the fact that B = P can then be used to infer that D - Q F Z does <u>not</u> represent PYLON since this coded word does <u>not</u> begin with a B - i.e. the B can be (indirectly) relevant.

The most satisfactory way of obtaining items with additional data of both sorts - relevant and irrelevant - is to work from items with irrelevant letters all the same, e.g.

(1)CLOUD (2)COUNT (3)CIDER (4)CHAIN (5)CARTS

X	-	J	-	P	....	( )
-	Q	-	P	G	....	( )
X	J	Z	G	-	....	( )
X	-	F	-	Z	....	( )
X	-	Q	-	F	....	( )

<u>Item VI</u>

Included here are the J's, P's and Z's of Item I; the additional relevant letters (Q's, G's and F's) of Item II; and the additional irrelevant letters (X's) of Item III. The items have been so devised that, when both additional relevant and additional irrelevant letters are present, no coded word is ever completely filled in: they all have at least one blank space. To facilitate comparison, Items I, II and III are set out together below.

(1)CLOUD (2)COUNT (3)CIDER (4)CHAIN (5)CARTS

Item I						Item II						Item III				
-	-	J	-	P		-	-	J	-	P		X	-	J	-	P
-	-	-	P	-		-	Q	-	P	G		-	-	-	P	-
-	J	Z	-	-		-	J	Z	G	-		X	J	Z	-	-
-	-	-	-	Z		-	-	F	-	Z		X	-	-	-	Z
-	-	-	-	-		-	-	Q	-	F		X	-	-	-	-
Just sufficient data						Additional relevant (redundant) letters						Additional irrelevant letters				

At one stage, when items with irrelevant letters all different were being used, it was thought that items with the coded words completely filled in might be regarded as containing both redundant and irrelevant data, e.g.

(1)FOLKS (2)TABLE (3)DRIFT (4)LIKEN (5)STEAK

P	G	H	C	J	..............( )
J	R	F	Q	M	..............( )
Q	H	N	M	Z	..............( )
C	V	Q	N	U	..............( )
U	J	M	R	N	..............( )

Item VII

Item VII includes all the letters of Item IV, plus other letters that are, for the most part, redundant. Scrutiny reveals that the letters in the second position, deemed irrelevant in Item IV, have become relevant now that all the coded letters are given. Thus the R of the second column is duplicated in the fifth coded word; the H in the first coded word; and the J in both the first and the second coded words. The V and the G are nowhere duplicated, and hence still not of any direct use. There is however the possibility of their being used negatively. Items with full coded data are therefore riddled with redundancy, to the almost complete exclusion of irrelevancy, and are not suitable as examples of problem situations containing both additional relevant and additional irrelevant data.

Items with full data have other drawbacks as well from the research point of view. While there is only one ordering of the uncoded words completely consistent with the pattern of coded letters, there is any number of ways of arriving at this ordering. A multiplicity of starting points exists, and there is scope for a variety of methods of attack. That marked differences in method do occur when full coded data are given has been noted by Donaldson (1956). Admittedly there are various possible starting points with less than full data too, but their number is limited and precisely known. And with less than full data the methods of attack that can be used are circumscribed by the particular letter pairs given. One method of attack that is ruled out is that of working from the uncoded array to the coded, rather than vice versa. Suppose that C V Q N U and U J M R N have, in Item VII, been identified as FOLKS and STEAK respectively. Since R = A the word TABLE can now be linked up with its coded counterpart merely by finding a coded word with an R in second place. In like manner, since Q = L the coded counterpart of LIKEN can be found simply by finding a coded word beginning with a Q. To proceed in this way when the coded words are less than complete is risky in the extreme. The positions about which predictions have been made will in all likelihood be blank. The tendency to work from one extreme of the data or the other (or to start from both extremes and converge upon the middle) is an issue of considerable interest. It has been discussed in terms of working forwards as opposed to backwards, from what is given as opposed to what is required (Duncker, 1945; Polya, 1957; Newell, Shaw & Simon, 1958, 1962; Raaheim, 1960). It is an issue that is not at all new, a very lucid account and assessment of the method of working backwards from what is required being quoted by Polya (p. 141 ff.) from the writings of Pappus, a Greek mathematician of around 300 A.D. But to give code items with full data, thus allowing the subject the option of working from either the coded or the uncoded list of words, only complicates an already complex situation unnecessarily. For these various reasons items with full coded data were abandoned early on in the research. They are only mentioned here because they <u>were</u> used at first, and because they happen to be the original of item type. They come from the Moray House Tests of Verbal Reasoning, although Greek letters are there used for the coded words and the number of words per item and the length of these words varies from item to item. The Moray House item type has been taken over and altered, somewhat radically, to suit the purposes of the present research.

Coding questions of one sort or another have long been in use as intelligence test items, but usually with a "rational" code based on some rule of alphabetical sequence, e.g.

> If in a certain code BNLA means COMB, what does EHDKC mean ?

Apart from Donaldson (1956) no research seems to have been done on the type of code item used here, and little has until now been known as to how they function. Since solution is dependent on structural rather than semantic features of the words, the items are more nonverbal than verbal in nature. It is always easier to specify and control the stimulus determinants of response for nonverbal than for verbal items, and nonverbal materials are often preferred to verbal for this reason (e.g. Kessen & Kuhlman, 1962, p.62). Various schemes have been proposed from time to time for constructing nonverbal items on the basis of systematic principles (e.g. Penrose & Raven, 1936; Burt, 1950; Menger, 1953; Dressel, 1957), and the present work is in this same tradition.

With respect to the initial choice of uncoded words, any word with doubled letters (as ANNOY) or with the same letter first and last (as LABEL) was excluded from consideration. So too, as far as possible, were proper nouns (as SPAIN); words with strong affective overtones (as DARKY); and little known words of esoteric meaning (as ALGID). Such words could not however always be avoided due to the at times very stringent structural demands of a set of five words as a whole. As for the relationship between uncoded letters and their coded counterparts, there was no requirement either of alphabetical proximity or of its opposite. The coded counterpart of a C might for example be a D, but it might just as easily be a Z. In order to minimise possible confusion between coded and uncoded letters, no letter from any of the uncoded words was used in the coded array for that item. It was sought also to minimise possible perceptual confusions among the different <u>coded</u> letters. While the perceptibility of the letters of the alphabet has been studied with the letters taken one by one (e.g. Enticknap, 1957) comparisons between pairs of letters seem not to have been made. And the perceptibility of the single letters has in any case been assessed in terms of threshhold measurements, something alien to the present concern, and the results vary somewhat according to the type face used. In the absence of empirical knowledge of the relative discriminability of different letters one can only proceed intuitively. It was considered that letters would be the more readily discriminable the more they differ as "shapes" - that, for example, the X's, J's and Q's of Item I below are more readily discriminable, each from the other, than are the X's, K's and Z's of Item I'.

```
(1)GROAN   (2)EVICT   (3)WEAVE   (4)ACRID   (5)TOUGH

X - - - -                        X - - - -
- - - - -                        - - - - -
- - - J X                        - - - K X
- Q - - -                        - Z - - -
- J Q - -                        - K Z - -

  Item I                           Item I'
```

Since it was not intended that the difficulty of the code items should lie in the task of perceptual search, a mixture of "rounded" letters (B's, Q's, etc.) and "angular" letters (X's, T's, etc.) was used wherever possible, rather than either kind alone.

Which letter pairs appear in the coded stimulus array is to some extent arbitrary. They must however be such as to yield an unequivocal solution, and be appropriately distributed among the five separate words (see pp. 14, 15, 21). It was further decreed that the two members of a letter pair should never both be in the same ordinal position. Thus Item I below would be admissible, but Item I' would not.

```
(1)FOUND   (2)BELOW   (3)CLASP   (4)HEATS   (5)CROWD

- - - M -                        X - - - M
- X - G -                        X G - - -
- - - - -                        - - - - M
- - - - G                        - - - - -
- - X - M                        - - G - -

  Item I                           Item I'
```

Letter pairs whose members are in the same ordinal position are easier to work with than are letter pairs with members in different ordinal positions (see Study 7), and indiscriminate mixing of the two sorts was therefore considered undesirable.

It may be noted in conclusion that the principles of item construction outlined above apply in full only to the items used in the final stages of the research. These items were the result of a long and gradual process of development, involving progressive refinement of the item type as more features of the items were brought under control. The items used in the preliminary studies did not measure up to all the criteria stated above.

****

III. A SURVEY OF THE LITERATURE PERTAINING TO THE STUDY OF REDUNDANT AND IRRELEVANT DATA IN PROBLEM SOLVING

A major reason for reviewing the literature on a given topic is that it sets the experiment(s) at hand in an appropriate context, and furnishes hypotheses for careful experimental testing. But due to the tendency in problem solving studies to present the subject with no more materials than sufficient for solution, there is no convenient background body of literature for the code item research. There are a few pertinent studies, cited below, but none of them uses materials strictly comparable to the code items. They therefore yield not so much "hypotheses for careful experimental testing" as "hypotheses that might prove fruitful if adapted to the code item situation". There are in addition various studies of tangential relevance to the code item work, and they too are mentioned below. They too are suggestive of "hypotheses that might prove fruitful if adapted to the code item situation".

(a) Problem solving studies having data additional to those sufficient for solution

The most pertinent of the previous studies are those on the effects of irrelevant data on concept attainment. Reed (1946) had 42 cards with a nonsense syllable on the back and from 4 to 12 words on front. One word belongs to the category represented by the nonsense syllable: the rest are irrelevant. The greater the number of irrelevant words the longer the time taken to learn the concepts. If the irrelevant words from one card to the next are inter-related, and so suggest alternative but incorrect categorisations, their effect is even more disturbing. In another experiment, by Johnson and Hall (1961), verbal concept problems were used thus:

bark	1. birch
salt	2. number
pepper	3. wood
cement	4. pot
paprika	5. cloves
thyme	

Four of the left hand words go together. The subject has to find a word from the same class among the words 1-5. The number of relevant and irrelevant words in the left hand column varied from 2 to 8 and from 0 to 6 respectively. The greater the number of irrelevant words, the longer the time taken to solve a problem.

Outwith the field of concept attainment, Bartlett (1958) in his work on thinking within "closed systems" gave his subjects a card on which were the words:

A, GATE, NO, I, DUTY, IN, CAT, BO, EAR,
O, TRAVEL, ERASE, BOTH, GET, HO, FATE.

E R A S E
F A T E

Words were to be chosen from those at the top of the card so as to complete the vertical arrangement indicated by ERASE and FATE, taking E R A S E as the middle word of the column. Subjects were told that "Not all the words given need be used". Proffered solutions were variable in the extreme, and only about 2% of them were of the form designated as "most fitting", viz.

A
B O
C A T
D U T Y
E R A S E
F A T E
G E T
H O
I

Taking this solution as the criterion, the words not used are "irrelevant". Their role is not explicitly discussed by Bartlett, but it would seem to be that of suggesting to the subjects alternative solutions that, to them, are fitting and satisfactory.

Experiments on fixation of method provide further instances of stimulus materials more than sufficient for solution, although the superfluous data are only present in order that Einstellung effects may occur. In Luchins' (1942) basic fixation-of-method experiment three measuring jars were specified, plus a required amount of water, e.g.

Problem	Jars regarded as given			Amount of water required
	A	B	C	
(i)	21	127	3	100 quarts
(ii)	9	42	6	21 quarts
(iii)	23	49	3	20 quarts

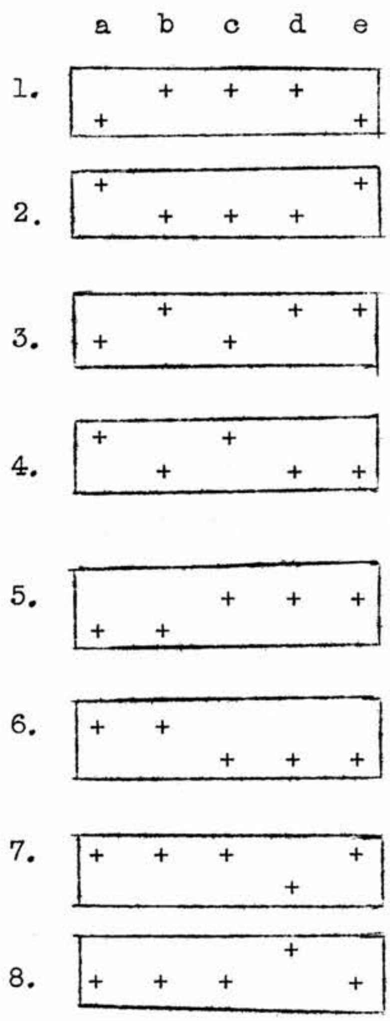
The formula B-A-2C yields the required amount of water in all three cases. Problem (iii) is however more simply solved by ignoring the largest jar. The data are here "redundant": there is more than one way of arriving at solution, and not all the data need be used. Luchins found that subjects right up to the post graduate level became so "set" for the B-A-2C procedure as to be blind to the opportunity, when it arose, of using a more direct 2-jar procedure. This might be seen as the development of so strong a set towards using all the data provided that the possibility of dispensing, in certain circumstances, with part (s) thereof is not even considered. Mention has already been made (p. 1) of the attitude that if something is given it is there to be used, or else it would not have been given in the first place. The extremes to which this attitude can be taken are quite extraordinary. 30% of a college class, asked to obtain 3 quarts from jars of 3, 65, and 29 quart capacities, were completely blind to the "obvious" solution of filling the 3 quart jar just once. They proceeded to the solution of 3 = 65 - 29 - 11x3. 62% of another college class, given the problem of arriving at 4 quarts of water from a 4, a 67 and a 17 quart jar, failed to achieve <u>any</u> solution in the 2½ minutes allowed due to their trying to use all three jars. It should however also be noted that control groups to which Luchins gave just an initial 2-jar problem followed by problems like (iii) above usually continued to use 2-jar procedures. This suggests that a set to use less than all the data may be established just as readily as is the set that all the data are to be used. What is <u>not</u> readily achievable is to induce subjects to adopt a particular one of these sets once they have come to accept the other.

As a variant on the standard water jar experiment a fourth and irrelevant jar was specified (Luchins & Luchins, 1950). In each case this fourth jar does not and cannot enter into the calculation of the amount of water required - or at least not in any easily discoverable way. Having a fourth jar means that the subject has to be selective of data. If for example the B-A-2C method is to be used then one, only one, and the proper one of the jars must be disregarded. Confronted with this demand for active search and selection of data the subjects were far more flexible, and less prone to Einstellung effects. But the extra irrelevant data also had a deleterious effect in that more total failures occurred than before.

Extra irrelevant data are also present in experiments on functional fixedness. Duncker (1945) used five "practical" problems in which objects crucial to solution were set amid a great jumble of objects on a table. Maier (1930, 1931, 1933), Bulbrook (1932), Székely (1950) and Saugstad (1955) have also used practical problems where one or more objects relevant to the task at hand are to be selected from a large number of other objects. The objects irrelevant to solution are found to obscure the crucial objects to the extent that the two are physically or functionally similar. If however a crucial object is somehow made conspicuous or "prägnant" (Duncker's term), the impact of the various irrelevant objects is then slight. That "seduction by the irrelevant" is difficult to resist when relevant and irrelevant stimuli are qualitatively alike, but not when they are distinctively different, is a very general finding. Luchins and Luchins for example found that if the irrelevant jar in their 4-jar problems was distinctive due, say, to its size, it was screened out without a great deal of trouble. If irrelevant data are <u>sufficiently obviously</u> irrelevant, they may not hinder solution at all. Thus Gottschaldt (1933), in an experiment quoted by Woodworth and Schlosberg (1954), had a short stick placed inside a child's play pen. Outside were two sticks, a long and a medium one, but only the long one would reach a "lure" beyond. The normal six year old disregarded the useless medium-sized stick completely, using the short one to pull in the long one which could then be used to obtain the reward. This issue of the "obviousness" of stimuli, and of differences between stimuli, turns up in a variety of contexts and under a variety of names: "conspicuousness", "cue salience", "discriminability of attributes", "formal similarity", "field homogeneity", etc. It raises a host of problems as to how the "obviousness" of a stimulus is to be measured, and by how much and in what way two stimuli must differ before the difference is an obvious difference. It is, unfortunately, all too true that "Surprisingly little is known about the factors that lead to confusion between visual objects" (Mackworth & Mackworth, 1958, p. 219).

The experiments of Bruner, Goodnow and Austin (1956) on selection strategies in concept attainment provide another instance of stimulus materials in excess of what are sufficient for solution. While their stimulus materials are not anything like the code items, their results are of interest and possible relevance. The subject is presented with an array of 81 cards on which there are figures differing in shape (square, circle, or cross), colour (red, green, or black) and number (single, double, or triple). Each card has one, two or three borders. Suppose the concept to be attained is that of "red figures". A positive instance (e.g. two red squares with one border) is pointed out to the subject, and he is to discover what the concept is by choosing other cards for testing. Any card with neither two figures nor red figures nor squares nor one border will in this case be irrelevant. Others of the cards will be redundant, which ones depending on how the subject proceeds. Many subjects were found to have a "thirst for confirming redundancy", and the helpfulness of redundant instances is discussed in some detail.

That redundancy will "help" the subject is a common assumption. Overcueing (having extra cues that overdetermine solution) is regarded as helpful in the field of programmed learning for example, especially by those of a Skinnerian bent (e.g. Skinner, 1963). But redundancy is by no means always an aid. Take Bricker's (1955) experiment on the identification of redundant stimulus patterns. He used the eight patterns diagrammed below:



Each + represents a light

These redundant patterns were compared with nonredundant patterns consisting of lights c, d and e only, they being sufficient to make each pattern distinctively different from every other. Redundancy had a deleterious effect on pattern identification: it only made perceptual discrimination the more difficult. The sort of redundancy at issue here is defined in terms of information theoretical notions, and is somewhat different from the redundancy of the code items. Bricker's experiment does however serve to demonstrate that redundancy is not invariably a help. This is implied as well in Underwood's (1952, p. 215) general theoretical proposition (not based on any solid empirical evidence) that "The greater the number of stimuli involved, the slower will be the rate of solution (of a problem). The stimuli may be relevant, irrelevant, or both, and the proposition still holds".

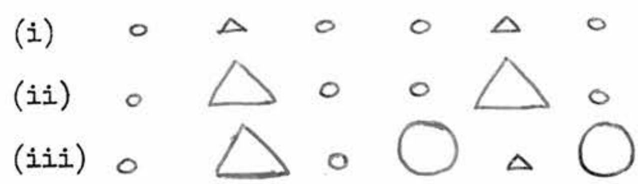
In some problem solving studies extra data, intended as a help, are systematically introduced. Speakman (1954) for example presented subjects who had failed to identify the values of different coloured stamps with a "clue card", supposedly helpful. The subjects were from 20 to 79 years old. For the younger subjects the clue card was a help, but not for the older subjects. Many of these regarded it simply as another complicating factor in a problem situation already quite complex enough. That supposed helps do not always function as such is noted also by Duncker (1945) and Wertheimer (1945). Among the problems used by Wertheimer in his work with schoolchildren was that of finding the area of a parallelogram. When various auxiliary lines were introduced, intended as helps, they were sometimes no help at all. Wertheimer (p. 61) concludes that the child who does not see the auxiliary lines in their role or function may receive them merely as "added complications" or "ununderstandable additions" that make the situation even more puzzling than before.

While the redundant data in these various problem situations hinder rather than help solution, they would seem to hinder not because "redundant" so much as because "additional". We human beings are not particularly adept at coping with large numbers of stimuli presented simultaneously. Burt (1919, 1921) has found with children that only the more mature can deal with large amounts of data all at once, and we become progressively less adept at so doing with age (e.g. Clay, 1954; Welford, 1958). Visual search tasks can present a large amount of data simultaneously, and such tasks are known to become more difficult with an increase in "display load" whether the added data are relevant or not. This is true for displays that are static (as for the code items) as well as for those that are changing over time (e.g. Mackworth, 1949; Conrad, 1955; Eriksen, 1955; Green & Anderson, 1956; Mackworth & Mackworth, 1958). Perhaps there is a point, for any stimulus array, beyond which extra data, relevant or irrelevant, will create a perceptual "overload". Nor will this overload be perceptual only: with a complex visual display one has to keep track of which parts have already been looked at and whether or not they are relevant, and this involves a strain on immediate memory. The time scale of immediate memory is very short, and the material held in memory very vulnerable to distortion or decay (e.g. Kay, 1953; Conrad, 1957, 1958; Brown, 1958; Broadbent, 1962). Subjects engaged in problem solving frequently complain of not being able to hold all the requisite data in mind. We as human beings operate at all times within ranges or spans, temporal and spatial, which set definite limits to the amount of data with which we can effectively cope.

Since the code item research is to use test items as problem situations, it would be appropriate here to review previous studies within the field of mental testing on the effects of additional redundant and irrelevant data on item solution. But test items with data over and above those logically sufficient for solution are, as has been pointed out, rare; and studies on the effects of such additional data are virtually nonexistent. Cattell (1940) has commented on the fact that many of Raven's matrices are overdetermined, and the matrices in his "culture fair" intelligence tests have some of the cells left blank. He has not however compared redundant with nonredundant matrices, or explored the role that additional redundant data might play. As for additional irrelevant data, Rimland (1959) has found that arithmetical reasoning items are significantly more difficult if extraneous numerical information is included. Why this should be so is not however considered, the purpose of the study being rather to determine whether including extraneous numerical information in an arithmetical reasoning test would lower its correlation with a verbal intelligence test without appreciably lowering its internal consistency. Buswell (1956) has also put irrelevant numerical information into arithmetical reasoning items, but his irrelevant data are not always additional to what are sufficient for solution, i.e. some of the items have data less than sufficient for solution. The irrelevant information caused considerable difficulty: there was no ready basis for discriminating between what was relevant and what not.

In conclusion let it be noted that in all the situations discussed above the extra data at issue have been stimulus "units" or "elements". As such they lead to an increase in the sheer amount of data present in a problem situation. There is on the other hand the possibility of "aspects" rather than units of the data being redundant or irrelevant. Take for example the following:

> Underline the four figures that go together because they are alike in some way :-



In (i) the only possibility is to classify by shape, the four circles going together. In (ii) classification may again be by shape, but may also be by size or by colour. The same classification is arrived at whichever dimension is used, since all three are perfectly correlated. The solution is overdetermined, and aspects of the data redundant. The presence of redundant aspects does not, be it noted, increase the absolute amount of data. In (iii) the figures again differ with respect to shape, size and colour, but the size and colour differences are here irrelevant to classification: any attempt to use them as bases of classification leads to two classes each of three figures, instead of to a class of four. As with redundant aspects so too with irrelevant aspects there is no increase in the absolute amount of data.

The redundancy and irrelevancy of the code items are a matter of stimulus units, not aspects, and this should be kept in mind in any reading of the literature. A very large number of studies on redundancy and irrelevancy concern stimulus aspects rather than units, e.g. the concept attainment experiments on redundant and irrelevant stimulus dimensions, such as those by McGuigan (1958), Peterson (1962) and the various experiments within the framework of information theory (e.g. Archer, Bourne & Brown, 1955; Bourne & Pendleton, 1958; Gelfand, 1958; Bourne & Haygood, 1959, 1960, 1961, 1964). The experiments on redundant and irrelevant stimulus dimensions in discrimination and reversal learning are another case in point (e.g. Restle, 1955; Harlow, 1959; Hodge, 1959; Sutherland, 1959; Mackintosh, 1963). In general it is found that redundancy facilitates and irrelevancy impedes performance. The deleterious effects of irrelevancy decrease however with practice, and neither is redundancy always a help. Green and Anderson (1956) for example found that redundant aspects <u>impede</u> the task of visual search if the subject does not know that aspects of the data <u>are</u> redundant, or does not know which aspects they are. But since all the experiments just mentioned are concerned with stimulus aspects rather than units they are, while suggestive of hypotheses that might prove fruitful if adapted to the code item situation, of tangential relevance only to the code item research.

****

(b) Studies on redundancy and irrelevancy within the framework of information theory.

There are certain affinities, but certain discrepancies as well, between the concepts of redundancy and irrelevancy as applied to the code items and the information theory concepts of redundancy and noise. Owing to the widespread usage, in the current psychological literature, of information theory terminology, some comment is in order as to where the present research stands with regard to information theoretical notions.

The word "redundancy", as used in the psychological studies that draw their inspiration from information theory, has a multiplicity of shades of meaning, and one can only agree with Staniland (1960, p. 160) that the word is carrying "too many and too varied burdens". Basically, however, redundancy is a matter of constraints on theoretical freedom and a lack of independence between signals. Written English is a favourite example for illustrating the concept, with the single letter as the unit of analysis. The letters in a passage of written English are not independent, but follow one another according to known probabilities. Redundancy is measured by the extent to which the sequence of letters departs from being random. The contrast between material that is redundant and material that is random is brought out in the work of Fitts and his associates (1956) and Anderson and Leonard (1958) on visual pattern recognition; and also in the work of Aborn and Rubenstein (1952, 1954) and Miller (1958) on the recall of strings of letters. Miller used letter strings composed of G's, N's, S's and X's. Redundant strings were generated according to specified rules of transition, while random strings were generated from a table of random numbers. Strings used include the following:

Redundant	Random
S S X G	G N S X
S X S X G	X G S S N
S X X X X S G	X X G N S G G

The absolute amount of data (number of letters) in these contrasted sets of stimulus materials is in each case the same. This is true also for the redundant versus nonredundant materials used in most other studies within the information theory framework.* For the code items on the other hand this is not so. Redundant code items - as pointed out on p.38 - have more stimulus units than do the nonredundant: six coded letter pairs as opposed to three. Also, the distribution and positioning of letter pairs in the nonredundant code items is in no sense "random". Information theory redundancy further differs from the redundancy of the code items in that it concerns the relationship of the stimulus or stimulus pattern to some external source, and is not assessable in intra-stimulus terms. Suppose one is presented with a pattern of A's and B's thus:

A B A B A B A B A B

This pattern could have been generated at random ($P_A = P_B = 0.50$ at every stage). If thus generated the pattern is one out of a possible

---

* But not all. Bricker (1955) for example (see p. 34) created redundant stimulus patterns by adding stimulus units (lights) to his nonredundant patterns.

total of $2^{10} = 1024$ patterns, and it is not redundant in the information theory sense. The pattern might on the other hand have been generated according to the principle that each letter is always followed by the other, and never by itself. In this case the pattern would be highly redundant in the information theory sense: it is one out of a possible total of only two patterns, as compared with the possible 1024 patterns when no rule of sequence is employed. Lacking knowledge as to how the pattern was generated, it is impossible to say if it is redundant in the information theory sense or not. For the code items, on the other hand, the redundancy is "in" the stimulus array, and there is no call for reference to any implied or background source.* This is an advantage, since the assumption that the subjects are familiar with the appropriate background source is, in the laboratory situation, often unrealistic. This point is well made by Staniland (1960), who criticises the sort of redundant materials used by Fitts, Miller, etc. because they fail to bear the mark of their origin. Owing to the discrepancies between information theory redundancy and the redundancy of the code items it might have been preferable to have used some word other than "redundancy" in the latter connection, but no suitable alternative suggested itself. There are in addition definite <u>similarities</u> between the two concepts, as is apparent from such statements as that redundancy implies that "something more is said or written than is strictly necessary to convey the message"

---

* Cf. Attneave's (1954) concept of redundancy in visual forms as being a matter of features inherent in the forms themselves - symmetry, continuity of outline, etc.

(Cherry, 1961, p. 116); or that there are "more distinctive cues than are actually required" (Garner, 1962, p. 195).

The information theory concept of "noise" is most simply defined as any random distortion of the message. In the presence of noise (without redundancy also present) only a probability statement can be made as to what the message might actually have been. While there are affinities between noise and the irrelevant letters introduced to the code items in that these letters are "useless signals", the irrelevant letters are not random in their distribution or positioning; they do not destroy or distort parts of the data relevant to solution; and they do not render the solution less than fully determinate. A further characteristic of noise is that it originates from the communication channel, rather than from the message source. Noise would never, in telecommunication, be injected into a signal at its source, the whole point of communication engineering being to reduce noise to as low a level as possible. In its psychological applications, on the other hand, "noise" does often refer to disturbances introduced at the message source itself, e.g. the perturbation of visual patterns employed by Fitts to create noisy patterns from ones that are noise-free. The irrelevant code letters follow in this tradition, they being introduced to and forming part of the "message source". They do not however "perturb" (i.e. distort) the stimulus array, and this is a basic and important difference.

Psychological studies within the framework of information theory have found that the effects of redundancy (without noise also present) depend on the form of the redundancy and the nature of the task. In pattern perception its effect is generally deleterious:

speed of performance invariably declines, and accuracy declines if restrictive time limits are imposed and sometimes even without such time limits (e.g. Attneave, 1955; Anderson & Leonard, 1958; Fitts et al., 1956, 1957). In concept attainment on the other hand the effects of stimulus redundancy are generally facilitative, provided it is "relevant" stimulus redundancy (e.g. Bourne & Haygood, 1959, 1961, 1964). The effects of noise (without redundancy also present) are, for all tasks, generally adverse. In the majority of experiments both redundancy and noise are simultaneously present. Noise is destructive of the message and gives rise to errors, and redundancy in any shape or form here facilitates performance and plays an error-combatting role. This is true even for the sorts of redundancy that impede performance when no noise is present. But while the findings are here unequivocal they are scarcely germane to the code item research. The irrelevant letters in the code items not being destructive of the message (i.e. the relevant letter pairs), there is no call for redundant data to offset their effect. The bearing of the various studies within the framework of information theory on the code item research seems, at best, tangential.

It might in conclusion be noted that it was in order to avoid confusion with "information" in its narrow technical meaning (viz. the statistical rarity of signals from a specified source) that this thesis has been written up in terms of the amount and kind of "data" present in a problem and not the amount and kind of "information". The word "information" in its broad everyday meaning would have been just as appropriate, and definitely more manageable, but it might have been imbued for some readers at least with restrictive information-theoretical overtones.

(c) Redundancy and irrelevancy in relation to intelligence and intelligence testing.

Despite the lack of agreement, much commented on, between one definition of intelligence and the next, there has been repeated emphasis on an appreciation of what is relevant. Thurstone (1924) characterises intelligence as the capacity for abstracting the relevant aspects of a situation and inhibiting the irrelevant; and McKellar (1957, p. 101) argues likewise. Heim (1954, p. 29) considers that "intelligent activity consists in grasping the essentials in a given situation and responding appropriately to them"; while Porteous (1959) sees the intelligent person as one who is adept at selecting, from the various stimuli available, those most relevant to the situation at hand.

Tests of intelligence, as has frequently been observed (e.g. Wechsler, 1950) rarely assess what the definitions assert "intelligence" to be. Certainly, despite the various statements above and others like them, only a very few item types in use in tests of intelligence are specifically concerned with assessing an appreciation of relevance. Absurdities are one example. They were used as early as Binet, and are discussed by Piaget (1928) in connection with the development of intelligence in childhood. Then too there is the sort of item found in Burt's Tests of Graded Reasoning where, given a set of statements, one has to decide whether a certain conclusion is valid, or what conclusion can validly be drawn. But never has the obvious sort of item for assessing appreciation of what is relevant been exploited, viz. items having extra irrelevant data that are to be screened out.* True it is indeed that "The choice of generalisation, or the process of selection and rejection of information are very rarely demonstrated in intelligence tests" (Whitfield, 1951, p.197). That items with extra irrelevant data might prove useful as measures of intelligence is suggested by Buswell's (1956) finding of a positive relationship between ability to discriminate between relevant and irrelevant numerical information in problem arithmetic items and Thurstone P.M.A. scores - if P.M.A. scores can be taken as an adequate criterion measure of intelligence. There is however the danger that irrelevant data may suggest to subjects of superior intellectual ability plausible but incorrect hypotheses, not contemplated by the less intelligent. If this happens then it will be the less rather than the more intelligent subjects who score highest in the face of irrelevant data (see for example Osler & Trautman, 1961).

As well as the notion that the intelligent person is one who can cope with irrelevancy there is also the notion, espoused in particular by Bartlett (1951, 1958), that the intelligent person is one who can work from a bare minimum of cues, i.e. who can do

---

* Classification items, both paper-and-pencil and performance (i.e. sorting tests) are an exception: they almost invariably have stimulus aspects (dimensions) irrelevant to classification, i.e. aspects that vary in value from instance to instance but upon which classification is not based. In the Vigotsky Test for example both colour and shape are "non-relevant variables, which must be disregarded in arriving at the correct solution" (Semeonoff & Trist, 1958, p. 8). Such irrelevant data are integral to the item type. For classification items where class members are to be segregated from non-members the non-members might be regarded as irrelevant stimulus units. Again, such irrelevant data are integral to the item type - an item without them is unthinkable.

without redundancy. Bartlett has never put this idea to experimental test, but some support for it comes from McColskey's (1957) finding that the more intelligent child requires "significantly less structure" to succeed on a complex conceptual problem than does the child of average ability. Corman's (1957) finding that it is the less intelligent who profit most from extra redundant information might also be mentioned here, although the "information" he is concerned with is not stimulus information but verbal information about rules and methods. Then too there is Stevenson's (1962) assertion that a given response will be produced in subjects of higher intelligence by a smaller number of stimulus elements than needed for the same response to be produced in subjects of lower intelligence, although the "responses" he is talking about are conditioned responses.

The two lines of thinking mentioned above are brought together by Wohlwill (1962). He argues that the demand on intelligence will increase (i) with an increase in the amount of irrelevant data and (ii) with a decrease in the amount of redundant data present in a situation. Both predictions are confirmed as far as stimulus "aspects" are concerned.* If the same holds true for stimulus units then this would be a cogent reason for using the sort of code item described in Section II in tests of intelligence: they can be manipulated so as to have more or less irrelevant data and more or less redundant data as required. The choice of item types for inclusion in tests of intelligence has long been - and still is -

---

* Wohlwill's work cannot be regarded as a test of Bartlett's hypothesis that the more intelligent person is one who can work from a bare minimum of cues. Bartlett was concerned with stimulus units, not with stimulus aspects.

a vexed issue (e.g. Cook, Heim & Watts, 1963). Pleas for the construction of intelligence tests to be based on a coherent theoretical rationale (e.g. Donaldson, 1956, 1963; Loevinger, 1957; Furneaux, 1960) have had little impact on intelligence testing practice, and it is still in 1964 all too true that

> "Current intelligence tests are made up of a potpourri of tasks, assembled to some extent through historical accident" (Maccoby, 1964, p. 225)

and that

> "Large numbers of items are compiled on an intuitive basis, and the reasons for choice remain implicit or vague" (Smedslund, 1964, p. 252).

Of course, to assess the validity of the various versions of the code items as measures of intelligence requires an adequate criterion measure of intelligence. Scores on tests in current use are less than satisfactory in view of the fact that these tests do not test what intelligence is asserted to be, and often test no more than the speed with which one can solve trivial and seemingly senseless problems without error. But in the absence of any ultimate criterion of intelligence there is little alternative but to accept scores made on tests in current use. One can at least see whether or not those who score highest on traditional tests are also these who cope best with the irrelevant code items; or those who cope best with the nonredundant code items; or both.

****

(d) <u>Conclusions to be drawn from the survey of the literature pertaining to redundant and irrelevant data in problem solving</u>

<u>A</u>. Subjects can be expected to try and use all the data with which they are presented, redundant and irrelevant data included, unless they are specifically instructed otherwise.

<u>B</u>. Redundant data may

(i) aid solution by playing a confirmatory role, or

(ii) hinder solution by acting as extra complicating detail.

<u>C</u>. Redundant data, even if helpful, can be expected to lead to longer solution times due to the greater bulk of data to be worked through.

<u>D</u>. Irrelevant data can be expected to hinder solution, except when they are so "obviously irrelevant" as to be ignored from the start. The effect of irrelevant data will be the more disturbing

(i) the more irrelevant data there are

(ii) the less "obviously irrelevant" they are

(iii) the more they suggest alternative but incorrect solutions

(iv) the less practised the subjects are at the task.

<u>E</u>. The more intelligent subject may cope better in the face of irrelevant data, while the less intelligent may profit more from redundant data.

The most general conclusion to be drawn from the survey of the literature is that there is no general conclusion: the effects of both redundant (superfluous) and irrelevant (misleading) data vary with the tasks and the subjects used.

IV. REDUNDANT AND IRRELEVANT DATA AND THE SOLVING OF CODE ITEMS: AN ACCOUNT OF THE EMPIRICAL RESEARCH

It has been concluded from the survey of the literature that the effects on problem solving of redundant and irrelevant data will vary with the tasks and the subjects used. No previous studies have employed the sort of code items described in Section II, and rather than speculate in armchair fashion as to what effects redundant and irrelevant code letters might have it was preferred to see what effects in fact they do have, and then to seek an explanation of these effects. In consequence, no precise hypotheses were formulated at the outset of the research and the early studies were of a "let's look and see what happens" variety. From the results of these early studies, considered in the light of the relevant literature and in relation to subjects' comments, hypotheses were developed for experimental testing. These hypotheses were modified and refined as the research progressed, and resulted ultimately in the precise hypotheses tested in the final stages. The various studies are written up below in chronological sequence so that this process of development may be followed through.

(a) Studies involving group testing

Study 1

Three 10-item tests were constructed, using the same sets of words for each. In Test I the items have just sufficient data for solution; the items of Test II have extra irrelevant letters, all different; while in Test III the coded words are given in full, e.g.

(1)WRAITH (2)MARKET (3)HAWSER (4)FAVOUR (5)CINDER

	-	G	-	-	A	-	.............( )
<u>Item from Test I</u>	-	-	-	-	C	-	.............( )
	-	-	-	-	-	C	.............( )
(just sufficient data	-	-	-	-	A	-	.............( )
for solution)	-	G	-	-	-	-	.............( )

	F	G	-	Q	A	-	.............( )
<u>Item from Test II</u>	Z	-	-	H	C	-	.............( )
	W	-	-	I	-	C	.............( )
(extra irrelevant	Y	-	-	J	A	-	.............( )
letters)	O	G	-	M	-	-	.............( )

	F	G	Z	Q	A	X	.............( )
<u>Item from Test III</u>	Z	X	G	H	C	F	.............( )
	W	G	X	I	A	C	.............( )
(coded data	Y	H	V	J	A	X	.............( )
given in full)	O	G	T	M	B	X	.............( )

4 of the items in Test II had a single column of irrelevant letters, while the other 6 items had two columns of irrelevant letters (as in the item above). The number of words per item varied from four to seven while the length of the words varied from four letters to six letters, the rule of construction not yet having been developed that the number of words per item and the length of these words should be the same from item to item. Nor had the rules yet been developed

(i) that there should not be any words with double letters or with the same letter first and last

(ii) that no letter in any uncoded word should appear in the coded array for that item

(iii) that the two members of a coded letter pair should never both be in the same ordinal position

(iv) that the different coded letters should be as unlike in "shape" as possible

(v) that the coded letters deemed irrelevant to solution should not be negatively or indirectly relevant

(vi) that items with extra irrelevant letters should all have the same number of such letters

(vii) that the number of initially-plausible identifications per coded letter pair should be the same for all the letter pairs in any one item.

The subjects for Study 1, as for all the group testing studies, were First Ordinary Psychology students at the University of Edinburgh. Testing was done during scheduled tutorial meetings, the three tests being given to three different tutorial groups. The three groups did not differ significantly* in mean score on the AH5 Test of High-Grade Intelligence, and were assumed equivalent with respect to variables relevant to the solving of code items. Each test was prefaced by a single sample item, analogous in structure to the items in the test proper but with only one possible identification for each of the given coded letter pairs, i.e. no false leads (see Appendix I). Since subjects were likely to have had experience of "rational" codes, stress was placed on the fact that the coding principle for each item is arbitrary. It was also emphasised that the solution of each item is unambiguous, only one ordering of the uncoded words being completely consistent with the pattern of coded letters. A time limit of 15 minutes was imposed. The margins of the test sheets could be used for scribbling. The results are set out in Table 1.

---

* Any statement to the effect that a difference is not significant implies a P value $>$ 0.05

	Test I Just sufficient data (N=27)	Test II Extra irrelevant letters (N=25)	Test III Full coded data (N=26)
Number of whole items correct (max. score=10)	$\bar{X}$ = 6.0 R = (3-8)	$\bar{X}$ = 4.1 R = (2-8)	$\bar{X}$ = 5.9 R = (2-9)
Number of single words attempted (max. score =58)	$\bar{X}$ = 32.2 R = (16-48)	$\bar{X}$ = 24.6 R = (12-40)	$\bar{X}$ = 31.8 R = (19-48)
Number of single words correct (max. score =58)	$\bar{X}$ = 31.3 R = (16-46)	$\bar{X}$ = 21.8 R = (11-38)	$\bar{X}$ = 31.2 R = (17-48)
Number of errors (words incorrect)	$\bar{X}$ = 0.9 R = (0-4)	$\bar{X}$ = 3.0 R = (0-9)	$\bar{X}$ = 0.6 R = (0-5)
Error ratios = $\frac{\text{total words incorrect}}{\text{total words attempted}} \times \frac{100}{1}$	2.9%	11.3%	1.9%

Table 1 (R = range)

The over-all difference between groups is significant for all measures except number of errors (Median tests extended for k independent samples; all P values < 0.01, 2-tailed). Mann-Whitney U tests then reveal that the significant differences are between Tests I and II and Tests III and II in every case (all P values < 0.01, 2-tailed).* Hence it

* Nonparametric tests were used here and in subsequent studies because some of the score distributions were skewed, and it was preferred not to be bound by the restrictive assumptions underlying the corresponding parametric tests. Although the nonparametric tests used do not test differences in central tendency by comparison of means, mean scores are given in the tables of results due to their being the most appropriate summary statistic and most revealing of trends.
By "measures" is meant everything except error ratios.

may be concluded that irrelevant letters significantly impede solution of the code items as regards both speed (number of words attempted) and accuracy (number of errors). This effect cannot be attributed to the sheer number of coded letters to be worked through: the Test II items do have more coded letters than the items of Test I, but the Test III items with full coded data - and hence the most letters of all - were solved just as rapidly and no less accurately than those with just sufficient data, i.e. fewest letters of all. A perceptual factor may be involved: the extra irrelevant letters in the Test II items may partially obscure the relevant letter pairs and make them the more difficult to find, whereas with full data, where there are few irrelevant letters anyway (see p. 22) and where one can use whichever coded letter pairs one chooses, this ceases to be of importance. An attitudinal factor may also be operative: with full data it is readily appreciated that not all the coded letters have to be used, but if <u>not all</u> the coded letters are given then the subject may assume that those that <u>are</u> given are there to be used. (No statement was made in the instructions as to whether all the coded letters given should be used or not). One subject who did the items with extra irrelevant letters commented that "the tester would not provide more letters than are necessary - they (i.e. testers) are like that". Repeatedly is it found that people engaged in problem solving adopt false assumptions like this (e.g. in the "9 dots" and "4 matches" problems used in demonstrating fixations or sets, and see also Morgan, 1944).

Pearson product-moment correlations between number of coded words correct and AH5 scores are as follows (none of the score distributions markedly skewed):

For Test I items (just sufficient data)	r = 0.44
For Test II items (extra irrelevant letters)	r = 0.11
For Test III items (full coded data given)	r = 0.52

r = 0.11 does not differ significantly from zero, and the difference between r = 0.44 and r = 0.52 is not significant.

Splitting each group of subjects into those above versus those below the median AH5 score for that group, the mean number of coded words correct for each is:

	Test I Just sufficient data		Test II Extra irrelevant letters		Test III Full data	
	Above Mdn. (N=12)	Below Mdn. (N=12)	Above Mdn. (N=11)	Below Mdn. (N=12)	Above Mdn. (N=13)	Below Mdn. (N=12)
Mean number of coded words correct	34.0	28.3	22.7	22.9	33.9	28.8

Results for Tests I and III again stand together in opposition to those for Test II, although none of the intra-group differences between those scoring above versus those scoring below the group median AH5 score is significant. From these results it would seem that, at least for the sort of code item used here, intelligence as measured by the AH5 Test is not concerned with the ability to sift out what is relevant from what is not. This is interesting in view of Heim's (1954, p. 29) assertion that "intelligent activity consists in grasping the essentials in a given situation and responding appropriately to them". The results lend some support to the notion that the intelligent person is one who can work from a bare minimum of cues: those above the group median AH5 score fared better in the face of just sufficient data than did those below the group median. But it is also true that those above the group median AH5 score fared better in the face not of just sufficient data but of data highly redundant and overdetermined (Test III), and the correlation between AH5 scores and number of coded words correct is of the same order for items with full data as for the items with only just sufficient data for solution.

To summarise the findings of Study 1:

(i) items with extra irrelevant letters (all different) were found significantly more difficult than items with just sufficient data for solution and items with full coded data

(ii) items with just sufficient data for solution and items with full coded data did not differ significantly in difficulty

(iii) intelligence as measured by the AH5 Test of High-Grade Intelligence was found not to be concerned with the ability to sift out code letters that are relevant to solution from those that are not

(iv) limited support was found for the notion that the intelligent person is one who can work from a bare minimum of cues.

Study 2

Study 2 further explores the effects of additional data on the solving of code items. Four 10-item tests were constructed, with the same sets of words used in each. In the first the items have just sufficient data for solution; in the second there are extra irrelevant letters, all different (4 items with one column of irrelevant letters and 6 items with two columns); in the third the items have three redundant letter pairs (a new departure); in the fourth the coded words are given in full. Greater control was exercised over item structure than in Study 1. Each item consists of five 5-letter words, none of which has any doubled letters or the same letter first and last. No letter appearing in an uncoded word appears also in the coded array for that item (except for a few items with full data, where this could not be avoided), subjects in Study 1 having reported some confusion between coded and uncoded letters. All four tests were prefaced by the same two sample items, whose working was gone through in detail and aloud by the tester (see Appendix I). In the first example the coded words are given in full; in the second there are only just sufficient coded letters for solution. It was pointed out to the subjects that items with an intermediate number of coded letters would also be possible, but no example of an item of this sort was given, i.e. the subjects who did the items with extra irrelevant letters and with extra redundant letters had no direct experience thereof until the test proper. The tests were so distributed that every fourth member of a tutorial group received the same version. Subjects were told of this arrangement in case they should notice and worry about the fact of their own test being somewhat different from that of their neighbour. The four groups did not differ significantly on a preliminary code test (given prior to the testing proper, although scored subsequent thereto), and were considered "matched" for the purposes of the research. A 20 minute time limit was imposed. Results are set out in Table 2.

	Test I Just suffic. data (N=39)	Test II Extra irrel. letters (N=38)	Test III Extra redund. letters (N=39)	Test IV Full data (N=38)
Number whole items correct (max. score=10)	$\bar{X}$= 3.4 R=(2-10)	$\bar{X}$= 2.8 R=(2-6)	$\bar{X}$= 4.2 R=(1-9)	$\bar{X}$= 4.4 R=(0-10)
Number single words attempted (max. score=50)	$\bar{X}$=19.6 R=(15-50)	$\bar{X}$=17.2 R=(16-50)	$\bar{X}$=23.7 R=(20-50)	$\bar{X}$=24.8 R=(20-50)
Number single words correct (max. score=50)	$\bar{X}$=18.3 R=(15-50)	$\bar{X}$=15.3 R=(15-45)	$\bar{X}$=22.4 R=(14-47)	$\bar{X}$=23.1 R=(14-50)
Number of errors (single words incorrect)	$\bar{X}$= 1.4 R=(0-9)	$\bar{X}$= 2.0 R=(0-11)	$\bar{X}$= 1.2 R=(0-10)	$\bar{X}$= 1.2 R=(0-7)
Error ratios	6.8%	11.2%	5.1%	4.6%

Table 2

As in Study 1, the over-all difference between groups is significant for all measures except number of errors (Median tests extended for k independent samples; all P values $<$ 0.02, 2-tailed).

Mann-Whitney U tests then reveal significant differences between groups as follows:

Measure	Tests
Number whole items correct	II vs III* II vs IV *
Number single words attempted	II vs III* II vs IV *
Number single words correct	I vs III* I vs IV * II vs III** II vs IV **

(*= $P < 0.05$, 2-tailed; and **= $P < 0.01$, 2-tailed)

In contrast to Study 1, none of the differences between items with just sufficient data and those with extra irrelevant letters reach the 0.05 level of significance (Tests I vs II). No reason other than sampling fluctuation suggests itself for this: the items and subjects were comparable in each case, and while the subjects who did Test II in the present study were not given a sample item with irrelevant letters (as were the subjects in Study 1 who did the test with items having extra irrelevant letters), this would if anything cause the difference between items with just sufficient data and those with extra irrelevant letters to be greater rather than less than before. Also in contrast to Study 1 is the fact that items with just sufficient coded letters for solution are found more difficult than those with full coded data. That the items with extra redundant letters should be aligned with those having full coded data is not surprising: items with full data are, as pointed out on p. 22, highly redundant. (Items with extra redundant letters were not used in Study 1, so no comparison of results can be made). There was no opportunity in Study 2 to assess the bearing of intelligence (as conventionally measured) on performance on the code items, no intelligence test scores being available.

While few of the obtained differences between groups are significant statistically, the direction of the differences is quite clear: irrelevant data impede solution, redundant data facilitate it. These differences are not explicable in terms of sheer number of stimulus elements (cf. p. 54), a point worth making in view of the prevalent tendency for task complexity to be defined in terms of number of stimulus elements and a positive relationship then to be postulated between task complexity and task difficulty. If complexity equals number of stimulus elements and if complexity is a determinant of difficulty then, for the code tests, the positive relationship between complexity and difficulty does not hold: items with more coded letters are more difficult than those with fewer coded letters only if the extra letters are irrelevant to solution. Overemphasis on the quantitative features of stimulus materials is to be avoided; equally as important as how many stimulus elements are present is which particular elements they are, and the part they play in leading to solution. Bartlett's (1958) work on thinking lends support to this conclusion, and so too do such findings as that reaction time and memory span vary with the nature of the stimuli, independent of their number (e.g. Mowbray, 1960; Brebner & Gordon, 1964; Conrad, 1964; Stone & Callaway, 1964). The same sort of thing happens with the hearing of words presented dichotically (Dodwell, 1964). Still another indication that task difficulty does not always increase with an increase in amount of stimulus material comes from the finding (Osgood, 1957) that nonsense sequences that preserve the structure of English grammar (e.g. The maff vlems oothly um the glox nerfs) are easier to learn, despite the greater absolute amount of material, than are matched strings of nonsense items with the grammatical cues removed (viz. maff vlem ooth um glox nerf).

After the testing had been completed a questionnaire was distributed, the questions being as follows:

(1) In solving the code items did you consider all the given letters, or only some of them ?

(2) Did you formulate any rules of procedure that could be applied to each item in turn ?

(3) What role did 'trial and error' play in your attempts at solution ?

(4) How great a degree of confidence did you have in your obtained solutions ?

(5) Any other comments on the test items ?

The choice of questions was based largely on spontaneous comments made by subjects from Study 1. It was hoped that the questionnaire would throw further light on the ways in which subjects approach the different code items. The answering of the questionnaire was optional, and 90 copies in all were returned. Replies to question (1) were as follows:

	Percentage of Ss. stating that all letters used	Percentage of Ss. stating only some letters used
Items with just suffic. data (N=26)	69.2%	30.8%
Items with extra irrel. data (N=21)	19.2%	81.0%
Items with extra redund. data (N=25)	12.0%	88.0%
Items with full data (N=18)	0.0%	100.0%

(N's based on number of Ss. returning the questionnaire)

19.2% of subjects claim to have used <u>all</u> the letters in the items with irrelevant letters, although identification of the irrelevant letters can in no way advance solution. Some of the other 81.0% no doubt tried to use the irrelevant letters initially, but then realised their lack of worth. That only 12.0% of subjects made full use of the redundant letters is of interest: not much evidence here of a "thirst for confirming redundancy". This may be due to there being a time limit on the test: checking to see if the letter pairs as yet unused are consistent with the obtained solution is wasteful of precious time. That 30.8% of subjects given letter pairs "just sufficient" for solution claim not to have used them all seems indicative of limited vision and a lack of awareness of possibilities. Take for example the item:

(1)ROUGH (2)SLIDE (3)HUMID (4)MELTS (5)LATER

```
- - - J -    ..............(   )
- - J - Y    ..............(   )
C - - - -    ..............(   )
Y - - - C    ..............(   )
- - - - -    ..............(   )
```

Equating J with T and C with H would establish the first four coded words as MELTS, LATER, HUMID and ROUGH respectively, and the fifth coded word would then be SLIDE. This solution, arrived at using only the J's and the C's, happens to be "correct". But if only the J's and the C's are used there is also the possibility of equating J with I and C with R. This establishes the first four coded words as HUMID, SLIDE, ROUGH and LATER respectively, and the fifth coded word would then be MELTS. To ascertain which of these two solution is correct the third letter pair -

the Y's - must be brought into consideration.* That some subjects took only two letter pairs into account and did not go on to consider the third may again have been due to there being a restrictive time limit on the test. Whatever the reason, the questionnaire replies do suggest that subjects are not here so prone to wanting or trying to use all the data provided as the literature would lead one to expect. A subject who does not use all the given letters in an item with letters only just sufficient, logically, for a unique solution may be lucky and strike the correct solution straight off, unaware that alternative possibilities exist. He may on the other hand err. Errors made in this way will be of the sort that Donaldson (1963) calls "structural".** Apart from errors of this sort most other errors on the code items would seem to be "executive" - e.g. in the rush

---

* Cf. pp. 16-17

** A study of the errors made by children on a variety of problems of the sort used in verbal intelligence tests led Donaldson to distinguish three main error categories:

(i) structural errors, i.e. errors that arise from some failure to appreciate the relationships involved in the problem or to grasp some principle essential to its solution

(ii) arbitrary errors, i.e. errors that arise from a lack of loyalty to the given

(iii) executive errors, i.e. errors that arise not from any failure to understand how the problem should be tackled but from some failure in the actual carrying out of the required manipulations.

to beat the time limit a number may inadvertently be placed in the wrong pair of brackets. A clear instance of this comes from a test paper on which, for the item above, the subject had jotted down in the margin that C = H - which is correct - but had then placed the numbers 1 and 3, for ROUGH and HUMID respectively, in the pairs of brackets thus:

```
C - - - -      ..............( 1 )
Y - - - C      ..............( 3 )
```

The thing that emerges most clearly from the questionnaire is the extent to which individuals differ in their approach to one and the same item type. This comes out particularly in the replies to questions (2) and (3). A set procedure is outlined in connection with the sample items that preface the tests, and this procedure - if followed - will invariably yield the correct solution in the end. Some subjects, however, choose not to follow this procedure: they report it as being "too logical", and employ instead a more haphazard approach which relies very much on the perceptual "look" of the item and on intuitive feelings as to what any particular coded letter pair is likely to stand for.* Here is a further possible source of error, for an intuitive feeling that a particular solution "looks right" is bound at times to be fallible. To assert that large individual differences exist may well seem commonplace and trite, but it

---

* Cf. Jung's reference to two contrasted ways of making judgments:

(i) by thinking - a logical process, aimed at an impersonal outcome

(ii) by feeling - a subjective process, personal rather than impersonal does no harm to be reminded of the fact. It is all too easy to lapse into a "subtle personal parochialism" (McKellar, 1957, p. 19) and to commit the"egocentric error" (Hunter, 1963) of assuming that everybody else thinks in much the same way as oneself; and even while recognising that individual differences exist one may fail to recognise their enormous extent.

Despite individual differences there are preferences <u>shared</u> by the vast majority of subjects for some letter pairs rather than others. The most outstanding preference is for letters in the first and last positions. Take the item:

```
(1)PLANT   (2)TRACK   (3)UNTIE   (4)CLIMB   (5)GROUP

           D - - V -     ..............(   )
           V - J - -     ..............(   )
           - - - - D     ..............(   )
           - - - - -     ..............(   )
           - - - J -     ..............(   )
```

<u>Item I</u>

Almost all subjects - other things being equal - will begin with the D's. This is due in part to a D being at the beginning of the first coded word: the upper left hand corner of a stimulus array is known to dominate as an initial focus of attention (e.g. Lindley, 1897; Burack, 1950; Forgays, 1953; Bartlett, 1951, 1958; Berlyne, 1958). But this is not the whole explanation. If the coded words are re-ordered thus

```
V - J - -     ..............(   )
- - - - D     ..............(   )
- - - J -     ..............(   )
- - - - -     ..............(   )
D - - V -     ..............(   )
```

the vast majority of subjects will still begin with the D's. It is easier to scan the words (1)-(5) for words with the same letter first and last than for words with the same letter first and fourth (as for the V's) or third and fourth (the J's). The tendency to begin with letters in first and last positions dominates over the tendency to begin from the upper left hand corner. A typical reply to question (2) of the questionnaire is: "I started with words with first and last coded letters the same, rather than with middle letters alike, regardless of the positions of these words in the coded series". This preference for letters at the extremes rather than in the middle of words is intuitively understandable, and fits in with Hunter's (1959, p. 200) finding that in the solving of anagram problems "An extreme letter is, as it were, more manipulable than an intermediate one". Perceptual factors undoubtedly contribute to this: Bartlett (1951) has shown that the letters at the extremes of a horizontal list of eight letters are the more prominent to perception than are the letters in the middle.

A second but less pronounced and less general preference is for letter pairs with members in the same ordinal position, no matter what this may be. If Item I above were changed to

(1)PLANT (2)TRACK (3)UNTIE (4)CLIMB (5)GROUP

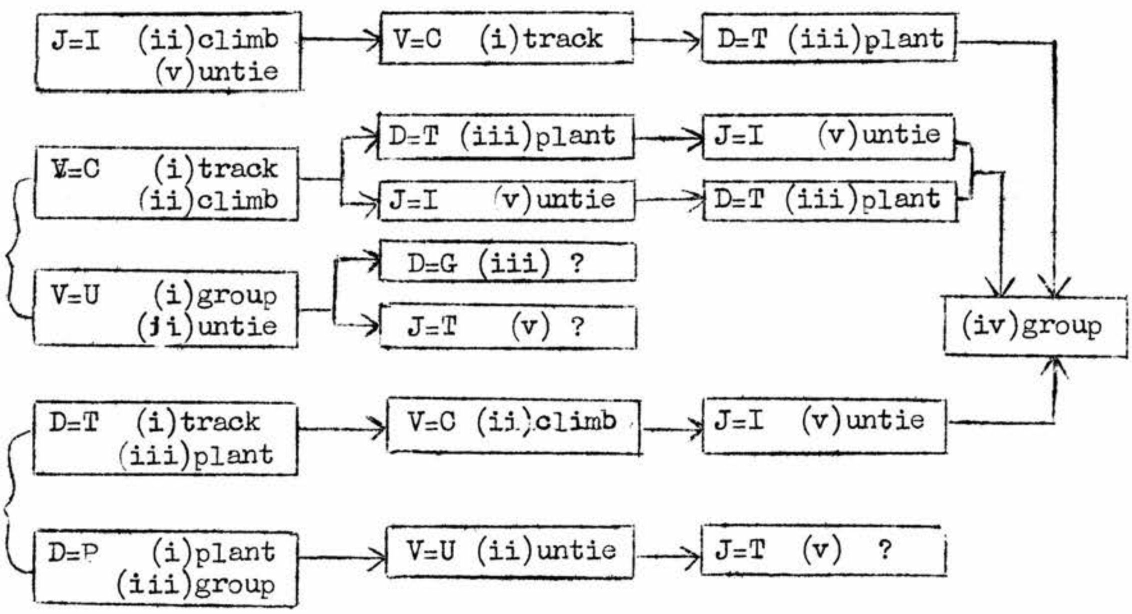
D	-	-	-	-	..............( )
-	H	J	-	-	..............( )
-	H	-	-	D	..............( )
-	-	-	-	-	..............( )
-	-	-	J	-	..............( )

Item II

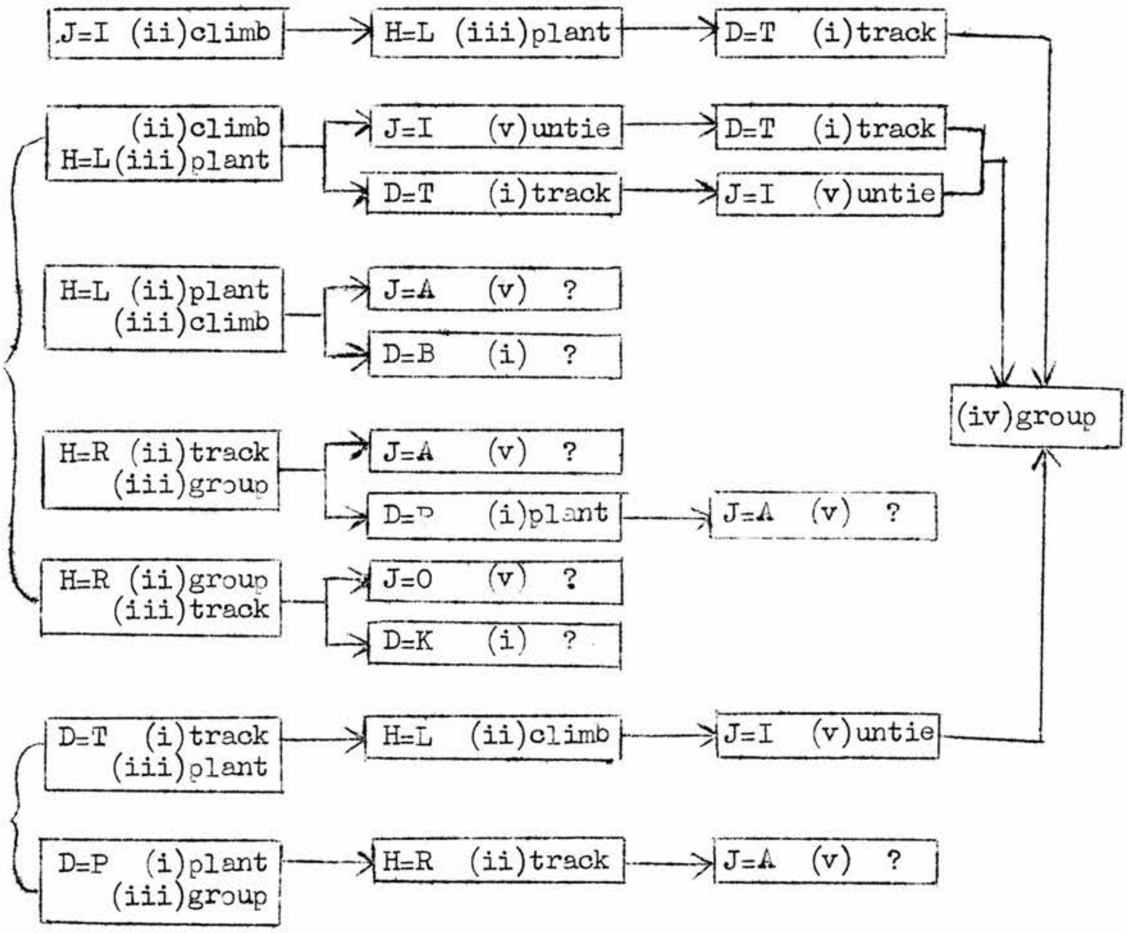
some subjects at least would begin with the H's. The tendency for letters in the middling positions to be avoided is here superseded. Should an item not have any coded letter pairs with members in the first and last positions or in the same ordinal position then the preference is for letter pairs with one member in an extreme position (first or last) or with members in adjacent ordinal positions. Thus patterns (i) and (ii) below are preferred to pattern (iii):

```
(i) - - - - X      (ii) - - - X -      (iii) - - - X -
    - - X - -           - - X - -            - X - - -
```

The fact that certain letter pairs have a low "preference value" does not mean that they are therefore of little use. Sometimes quite the opposite is true. In both Item I and Item II above the J's would have lowest preference value as far as their positioning is concerned, yet they are in each case the most opportune letter pair with which to begin. They open up no false leads at all, as the accompanying flow charts show.



Flow Chart for Item I



Flow Chart for Item II

53 of the 90 subjects who replied to the questionnaire had made no errors on the particular test they did. 26 of these 53 (49.1%) stated that they had 100% confidence in their obtained solutions, while another 23 (43.4%) stated that they had a high degree of confidence. Only 7.5% said their confidence in their obtained solutions was low. Taking now <u>all</u> the subjects (52) who expressed 100% confidence in their solutions, the number of errors made by these subjects ranged from zero to sixteen in a frequency distribution as follows:

Errors	Frequency
0	28
1	0
2	9
3	3
4	1
5	4
6	4
7	1
8	1
.	.
.	.
.	.
16	1

What these results indicate is that almost half of the subjects (43.4%) who did not err under-rated their performance, but not markedly except for 7.5%. On the other hand, many of the subjects who expressed 100% confidence were over-rating their performance, some of them to a very considerable degree.

To summarise the major findings of Study 2:

(i) the presence of letters irrelevant to solution - all different - was found to impede solution of the code items, although not to a statistically significant degree

(ii) the presence of redundant letters was found to facilitate solution of the code items, and to the same extent both for items with three extra (redundant) letter pairs and for items with full coded data

(iii) a majority of subjects claimed to have used all the coded letters both for the items with just sufficient data for solution <u>and</u> for those with extra irrelevant letters

(iv) a marked preference was reported for coded letter pairs with members in first and last positions in the words

(v) a second but less pronounced and less general preference was reported for letter pairs with members in the same ordinal position.

Study 3*

Study 3 is concerned with code items so constructed that their solution demands the use of indirect procedures. As an example take the following:

(1)FOLKS (2)TABLE (3)DRIFT (4)LIKEN (5)STEAK

-	-	-	C	-	..............( )
-	-	X	-	-	..............( )
Q	-	-	J	-	..............( )
C	-	Q	-	-	..............( )
-	-	-	-	-	..............( )

Item I

C occurs in first and fourth positions and could stand for F (the first coded word DRIFT and the fourth FOLKS) of L (the first coded word TABLE and the fourth LIKEN). Q occurs in first and third positions and can only stand for L. Hence the third coded word is LIKEN and the fourth FOLKS (and the first therefore DRIFT). Apart from the C's and the Q's the only other coded letters are a single X and a single J. Hence the coded counterparts of STEAK and TABLE cannot be identified by matching up pairs of coded letters with uncoded letters occurring in the same ordinal positions. Some alternative procedure must be adopted. And if

* The findings of this study are reported in an article entitled "On the solving of code items demanding the use of indirect procedures" and accepted for publication in the British Journal of Psychology.

the coded counterparts of STEAK and TABLE are to be identified other than by arbitrary allocation this can only be accomplished by a procedure that is indirect. It is known that C = F and Q = L and, since Q - - J - represents LIKEN, J = E. What X stands for is not known, but it does not stand for F or L or E (i.e. $X \neq F$ or L or E; or $X = \overline{F}$, $\overline{L}$ and $\overline{E}$). The fact that the second coded word has an X in third position can now be interpreted to mean that it cannot represent a word with an F or an L or an E in third position. This eliminates STEAK. Hence the second coded word is TABLE and the fifth STEAK, rather than vice versa. This final phase of the solution sequence is "indirect" in the sense (i) that it involves elimination of all (here only one) alternative possibilities by showing them not to be the case, and (ii) that it requires transformations of data from positive to negative form. Should items be desired in which there is a greater demand for indirect procedures, the coded letter pairs may be reduced to one, viz.

(1)FOLKS (2)TABLE (3)DRIFT (4)LIKEN (5)STEAK

-	Z	-	-	P	....	( )
-	-	X	-	-	....	( )
Q	-	-	J	-	....	( )
-	-	Q	G	-	....	( )
-	-	-	V	-	....	( )

Items are also possible whose solution is wholly indirect, no pairs of coded letters being given. It is however extremely difficult to construct such items within the framework of five 5-letter words, and no items of this sort were used in the research. It has been found by Bruner, Goodnow and Austin (1956) that subjects experience difficulty with indirect procedures involving transformations of data, and avoid them whenever possible. Bruner and his colleagues suggest that this effect may well be general to a variety of cognitive activities. Whether the avoidance of indirect procedures is due to inability to cope therewith or mere unwillingness to make the attempt is left an open question. Donaldson (1959) has shown the difficulty with and avoidance of indirect procedures to extend from categorisation to "matching" problems, and she claims unambiguous evidence of inability to cope therewith. This interpretation of her results can however be challenged: rather would it seem that while subjects may not be able competently to handle indirect procedures straight off, the ability to do so develops with a little practice.* Study 3 seeks to determine whether difficulty with and avoidance of indirect procedures extend also to the code items. If so, (i) what is the locus of the difficulty ? and (ii) is it inability to cope with the procedures, or unwillingness to make the attempt, that leads to the avoidance ? Should the "indirect" code item prove a useful item type the intention was then to ascertain the effects of redundant and irrelevant data on the solving thereof.

A 10-item test was constructed, using the same sets of words as for Study 2. The first 5 items each have two coded letter pairs, the other 5 only one. The number of single coded letters varies from item to item (1 - 6). In each case the total number of coded letters is just sufficient for a unique and unambiguous solution. In the first investigation the test was prefaced by the same two sample items as used in Study 2. Neither of them makes any demand for indirect procedures. (If sample items involving indirect procedures had been given, and the test items not then

---

* See the British Journal of Psychology article, in press.

found unduly difficult, it would remain a matter of speculation whether they may have been unduly difficult had the sample items not been given, i.e. had subjects not been acquainted with the appropriate procedures and instructed in their use.) A 20 minute time limit was imposed. The results appear in Table 3.

	Indirect items (just suffic. data) N=40	Direct items (just suffic. data) N=39
Number of whole items correct (max. score=10)	$\bar{X}$ = 1.9 R = (0-6)	$\bar{X}$ = 3.4 R = (2-10)
Number of single words attempted (max. score=50)	$\bar{X}$ =15.9 R = (0-35)	$\bar{X}$ =19.6 R = (15-50)
Number of single words correct (max. score=50)	$\bar{X}$ = 12.9 R = (0-32)	$\bar{X}$ =18.3 R = (15-50)
Number of errors (words incorrect)	$\bar{X}$ = 3.1 R = (0-17)	$\bar{X}$ = 1.4 R = (0-9)
Error ratios	18.9%	6.8%

Table 3

Also included in Table 3, under the heading "Direct items", are the Study 2 results for items with just sufficient data. In them there were three coded letter pairs per item, and no demand on indirect procedures. The two groups of subjects do not differ significantly on the preliminary codes test (again used in Study 3) and can be considered "matched" for the purposes of this study. The difference between groups is significant for both number of whole items correct and number of errors (Mann-Whitney U tests; $P < 0.05$ and $P < 0.02$, 2-tailed, respectively). The differences for the other measures, while not significant, are all in the direction indicating inferior performance on indirect items as regards both speed (number of words attempted) and accuracy (number of errors). Thus it would seem that many subjects are not able <u>spontaneously</u> to apply indirect procedures to the solving of code items.

To determine whether the difficulty of indirect code items is due to some difficulty inherent in the indirect procedures themselves or merely to the subjects' lack of acquaintance with such procedures, further testing was undertaken in which the indirect items were prefaced by an additional sample item whose solution involves indirect procedures (see Appendix I). The solution of this item was gone through in detail, every effort being made to ensure that the data transformations involved were fully understood by the subjects. Conditions were otherwise exactly as before. A matched group of subjects did the "direct" version of the test. The results are set out in Table 4, overleaf. The difference between groups is significant for number of whole items and number of single words correct (Mann-Whitney U tests; $P < 0.05$, 2-tailed, in each case). The difference in number of errors is not this time significant. From these results it may be concluded that the difficulty of indirect items, in comparison to direct, is not reduced by instructing the subjects in the use of indirect procedures prior to their taking the test: the instruction had little impact, and the difficulty of the indirect items is just as pronounced as before, if not moreso.*

	Indirect items (just suffic. data) N=20	Direct items (just suffic. data) N=19
Number of whole items correct (max. score=10)	$\bar{X}$ = 2.1 R = (0-6)	$\bar{X}$ = 3.8 R = (0-8)
Number of single words attempted (max. score=50)	$\bar{X}$ =16.4 R = (5-32)	$\bar{X}$ =21.6 R = (10-42)
Number of single words correct (max. score=50)	$\bar{X}$ =12.8 R = (0- 30)	$\bar{X}$ =19.6 R = (0- 42)
Number of errors (words incorrect)	$\bar{X}$ = 3.7 R = (0-17)	$\bar{X}$ = 2.1 R = (0-9)
Error ratios	22.3%	9.5%

Table 4

It might be objected that a single sample item is scarcely sufficient to acquaint a subject with indirect procedures, but even allowing this there would still seem to be some difficulty associated with or inherent in the indirect procedures themselves. This is further

* Cf. the finding of Bruner et al. (1956) that despite detailed explanation of the nature of disjunctive concepts subjects still persist in using information in ways quite unsuited to disjunction. Cf. also the finding of Jeeves and Dienes (1964) that most subjects who failed their problems involving mathematical groups did so because they sought rules of a sequential nature, despite warnings against this in the instructions.

revealed by comparing the patterns of errors for the two sets of items. Errors on the direct items are comparatively few, and are so distributed that no one word in any item is wrongly identified more frequently than is any other. Errors on the indirect items are far more prevalent, and concentrated on words whose identification entails the use of indirect procedures. It would seem that subjects are brought to a halt by these words, and resort is had to guessing (arbitrary allocation). The fact that on the indirect items one must either use indirect procedures or fail (apart from chance successes) is in contrast to the tasks used by Bruner and by Donaldson, where success could always be achieved by direct procedures alone, although at high "cost" or the expenditure of excessive effort. Since none of the indirect items is completely soluble without some use of indirect procedures, the high failure rate on them suggests inability, rather than mere unwillingness, on the part of many subjects to employ indirect procedures. The possibility does however remain that failure was preferred to the labour that the indirect procedures entail. As Bartlett (1958, p.84) points out, subjects sometimes convey the impression that they cannot do anything with a problem when all that this really means is that they are not sufficiently interested to go to the trouble of working out what to do with it.

Turning now to the nature of the psychological difficulty associated with indirect procedures, the first thing to be noted is that these procedures entail transformations of data. If "thinking" is something that involves more than an immediate response to the immediate external environment, the items whose solution demands the use of indirect procedures demand more thinking than do the items that are soluble merely by the direct perceptual matching of coded with uncoded letter pairs.* Not only do the indirect procedures associated with the code items entail transformations of data, but these transformations of data are such as involve a change of logical subject. Take Item I above, repeated here for ease of reference:

(1)FOLKS	(2)TABLE	(3)DRIFT	(4)LIKEN	(5)STEAK	
-	-	-	C	-	..............( )
-	-	X	-	-	..............( )
H	-	-	J	-	..............( )
C	-	H	-	-	..............( )
-	-	-	-	-	..............( )

Solution of this item involves a transition from J = E, a statement about J, to $X \neq E$ - a statement about something other than J. (See p.71) The opposite transition, from E = J to $E \neq X$, is probably easier: both are statements about E. But this transition, while equivalent to the former from the point of view of formal logic, is not the one that subjects doing the code items actually attempt.

Donaldson (1956) has tentatively suggested that transformations

---

* While in solving the "direct" items most subjects rely exclusively on the procedure of perceptual matching of coded with uncoded letter pairs, two subjects spontaneously introduced to these items, wherever possible, <u>indirect</u> procedures based on negative instances. Both subjects were B.Sc. students, which may be significant in view of the various indications in the literature (e.g. Chant, 1933; Dale, 1935; Burt, 1949; Wallace, 1952; van de Geer, 1957; John, 1957; Cook, Heim & Watts, 1963) that arts as opposed to science students, whether due to academic training or to earlier life experiences, have characteristically different modes of approach to problems.

of data are the more difficult if they involve a change of logical subject, and the present results substantiate this. Donaldson has further suggested (1959) that a "feeling of finality" may attach to positive statements, and this suggestion is also relevant here. Such a feeling would contribute to failure with the seemingly simple transition from $J = E$ to $X \neq E$, the positive $J = E$ being regarded as an end in itself rather than as also a possible means to a further end within the framework of the problem.

A further factor contributing to the difficulty experienced with indirect procedures is that the subjects were lacking in appreciation of the fact that any instance is, simultaneously, a member of more than one class. All A's for example are, at the same time, $\overline{B}$'s and $\overline{C}$'s and ..... and $\overline{Z}$'s. Difficulty in viewing the same thing as simultaneously of more than one class has been demonstrated repeatedly by Piaget - e.g. the difficulties experienced by the young child in grasping that number is both cardinal and ordinal, or in regarding daffodils as belonging simultaneously to the subordinate class "daffodils" and the supraordinate class "flowers". It is not only children who experience this sort of difficulty, judging by the failure of so many First Ordinary Psychology students to appreciate that a letter of the alphabet - A for instance - is a member not only of the class of A's but also of the various supraordinate "equivalence" classes $\overline{B}$'s, $\overline{C}$'s, ..., $\overline{Z}$'s that are the complementaries of the classes of the letters of the alphabet other than A itself. Why does there exist this lack of competence in dealing with classes such as the class of $\overline{X}$'s ? Heidbreder (1945, 1947, 1948) argues that people prefer to work with cues that are directly perceptually apprehensible. The various $\overline{X}$'s tend not to bear any particular perceptual similarity one to the other, and this may lead to a reluctance to work with them. But the problem is less a matter of reluctance than one of failure even to conceive of such a class as that of $\overline{X}$'s, a class which, when united with the X's, exhausts the contents of the supraordinate class "letters of the alphabet". Linguistic factors may have something to do with this. Classes such as the class of $\overline{X}$'s are purely verbal, the only feature the various class members have in common being that they can all be labelled "$\overline{X}$". In at least one experiment (Bruner & Olver, 1963) linguistic convention has been found to be used but infrequently as a basis of classification. The label "$\overline{X}$" is not in any case one that comes immediately to hand as being applicable - say - to an "A", and Heidbreder has found again and again that subjects strike trouble when asked to work with stimuli for which they have no ready-made verbal labels. Then Brown and Lenneberg (1954) have found that ease of recognition of different stimuli varies with the availability of verbal labels for them.

To summarise the findings of Study 3:

(i) difficulty with and avoidance of indirect procedures were found to extend to the sort of code item used in this study

(ii) many subjects seemed truly unable, rather than merely unwilling, to employ the indirect procedures that solution of the various items demanded

(iii) the source of the difficulty seemed to lie in the demand for data transformations that involve a change of logical subject, and/or the need to work with equivalence classes that are negatively defined and purely verbal.

The effects of redundant and irrelevant data on the solving of items of the "indirect" type are taken up and discussed in Study 9.

Study 4

This, the last of the group testing studies, takes up from Study 2. The main findings of Study 2 were that redundant letters significantly facilitate solution of the code items while irrelevant letters impede solution, although not to a statistically significant degree. The next question is *why* these effects occur. Dealing first with the redundant items, scrutiny of those used in Study 2 reveals a marked tendency for the extra letter pairs introduced to these items to have fewer initially-plausible identifications than for the original letter pairs from the items with just sufficient data. Compare for example the following:

(1)FOLKS (2)TABLE (3)DRIFT (4)LIKEN (5)STEAK

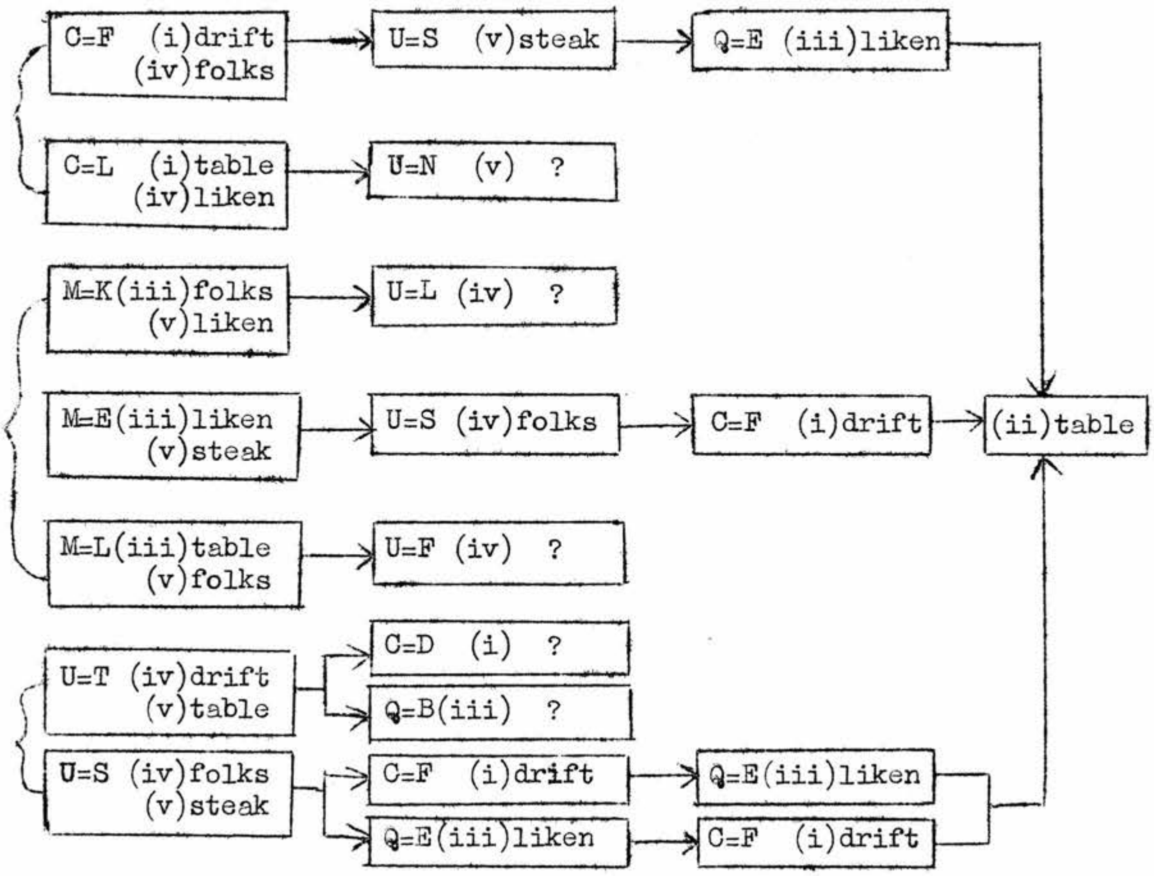
-	-	-	C	-
-	-	-	-	-
-	-	-	M	-
C	-	-	-	U
U	-	M	-	-

Just sufficient coded letters for solution

-	-	H	C	J
-	R	-	-	-
-	H	-	M	-
C	-	-	-	U
U	J	M	R	-

Three extra letter pairs

For the version with just sufficient coded letters there are 7 possible starting points, 4 of them false leads, as shown in the flow chart below:



Assuming for the moment that all three letter pairs are equiprobable as starting points, the probability of starting with any particular one of them (e.g. the C's) is 1/3. Given the C's as starting point there is then a probability of 1/2 of starting out on a wrong path. Hence the composite probability of starting out on a wrong path branching from the C's = 1/3 x 1/2 = 1/6. The composite probability of starting out on a wrong path branching from the M's = 1/3 x 2/3 = 2/9; while the composite probability of starting out on a wrong path branching from the U's = 1/3 x 1/2 = 1/6. Considering the item as a whole, the over-all probability of starting out on a wrong path = 1/6 + 2/9 + 1/6 = 10/18 = <u>0.56</u>. For the redundant version of the item there are the same 7 possible starting points as before, PLUS those opened up by the extra letter pairs (the H's, J's and R's). There is only one possible identification for each of these letter pairs (the correct one): H = L, J = T and R = A. This gives a total of 10 possible starting points, still only 4 of them false leads, and probabilities of starting out on a wrong path are now:

Letter pair	Probability of wrong path
C's	1/6 x 1/2 = 1/12
M's	1/6 x 2/3 = 2/18
U's	1/6 x 2/3 = 2/18
H's	1/6 x 0/1 = 0
J's	1/6 x 0/1 = 0
R's	1/6 x 0/1 = 0

The over-all probability of starting out on a wrong path is equal to the sum of these individual probability values = 11/36 = 0.31, i.e. the probability of starting out on a wrong path is less, for the redundant version of the item, than for the corresponding nonredundant version. For this particular item none of the extra letter pairs in the redundant version opens up any false leads. In other redundant items the extra letter pairs did open up false leads, but always fewer than for the original letter pairs from the nonredundant versions. In not one of the 10 items used in Study 2 was the probability of starting out on a wrong path greater, for a redundant item, than for the corresponding nonredundant item. The actual probability values were thus:

Item	Version with just sufficient data	Version with extra redundant data
1	0.37	0.37
2	0.61	0.56
3	0.67	0.44
4	0.50	0.25
5	0.33	0.33
6	0.33	0.17
7	0.50	0.36
8	0.25	0.21
9	0.44	0.22
10	0.37	0.36

These probability values are artificial to the extent that they rest on the assumption of all letter pairs being equiprobable as starting points. In actuality, certain letter pairs have higher "preference value" as starting points than do others (see pp. 65-7). But this may not be as serious a limitation as at first seems. Subjects often decide that a letter pair in a non-conspicuous and non-preferred position is likely to be the "best bet" as a starting point. This is particularly so for any item for which, on the item immediately prior, the letter pairs in positions of high preference value opened up an inordinate number of false leads. Some subjects may decide that letter pairs in these positions are worth trying again as starting points, but other subjects will decide the exact opposite. Human thinking is permeated with vagaries and idiosyncrasies, as has been frequently noted (e.g. Cohen, 1960; Herman & Engstrand, 1957), and there is no knowing for certain with which letter pair a subject will in fact begin any item.

In constructing the items for Study 2 the number of initially-plausible identifications per letter pair was not systematically taken into account. That the additional letter pairs in the redundant items opened up fewer false leads than the original letter pairs, thus reducing the over-all probability of starting out on a wrong path, happened quite by chance. That this was so does however suggest that the redundant items were easier than the nonredundant items because, by selecting appropriate subsets of letters with which to work, the subjects could relieve themselves of the burden of taking multiple alternative possibilities into account and of keeping track of the fate of them all. Many of the extra letter pairs introduced to the redundant items opened up no false leads at all, whereas most of the letter pairs in the nonredundant versions had at least two and often more initially-plausible identifications. There is evidence that a number of subjects set out deliberately to select those letter pairs that open up the fewest false leads, and even subjects who did not do this are likely, by chance, frequently to have chosen letter pairs in the redundant items that open up no false leads and hence, albeit unwittingly, to have made the task easier for themselves than it might otherwise have been.

The hypothesis that redundancy is helpful to the extent that it leads to a decrease in the probability of starting out on a wrong path is a hypothesis derived from the Study 2 results post facto. It must now be put to experimental test. In the present study the hypothesis is approached indirectly: if redundancy is helpful to the extent that it leads to a decrease in the probability of starting out on a wrong path then, if redundant items are constructed such that the redundant letters do <u>not</u> lead to a decrease in the probability of starting out on a wrong path, these redundant items will <u>not</u> be any easier than corresponding nonredundant items. Further, if the probability of starting out on a wrong path is markedly greater for redundant than for nonredundant items then the redundant items will be more difficult than the nonredundant, not less.

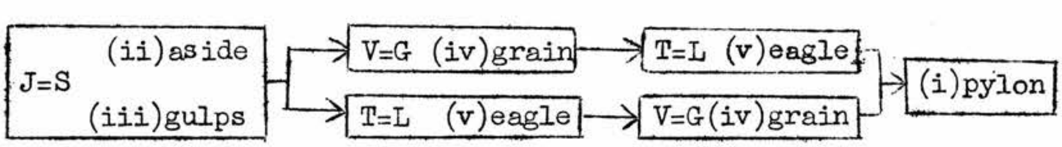
Three tests were constructed, each of 12 items. In Test I the items have just sufficient data for solution, and a certain letter pair is designated as the starting point, e.g.

```
(1)GRAIN   (2)ASIDE   (3)PYLON   (4)GULPS   (5)EAGLE

        -  -  -  -  -      ..............(    )
        -  J  -  -  -      ..............(    )
        V  -  T  -  J      ..............(    )
        V  -  -  -  -      ..............(    )
        -  -  -  T  -      ..............(    )
```

Item I

Subjects are told always to begin with the letters in the triangles. To encourage them to abide by this ruling, the triangled letter pair is always placed in a position of high "preference value". This usually means having the triangled letter pair towards the top of the stimulus array, and one member of the pair towards the top left hand corner. There is always only one possible identification for the triangled letter pair and, by taking it as the starting point, no false leads present themselves and solution is quite straightforward. For Item I it runs as follows:



(Note: For some of the items with just sufficient data and no false leads that data are, strictly speaking, more than sufficient for solution, e.g.

```
(1)FROST   (2)DRIFT   (3)AFIRE   (4)TARTS   (5)RADIO

                  - - L - U    ..............(   )
                  - K - - -    ..............(   )
                  - - - - L    ..............(   )
                  - - - - -    ..............(   )
                  - - K U -    ..............(   )
```

The L's establish the first coded word as FROST and the third as RADIO. If one then takes the K's, they can only stand for R. This makes the second coded word DRIFT and the fifth TARTS, and the fourth must therefore be AFIRE. Although solution can here be validly attained from two rather than three letter pairs, three pairs are always given in these items for the sake of uniformity with the items where there <u>are</u> false leads and where three letter pairs <u>are</u> needed to determine a unique solution).

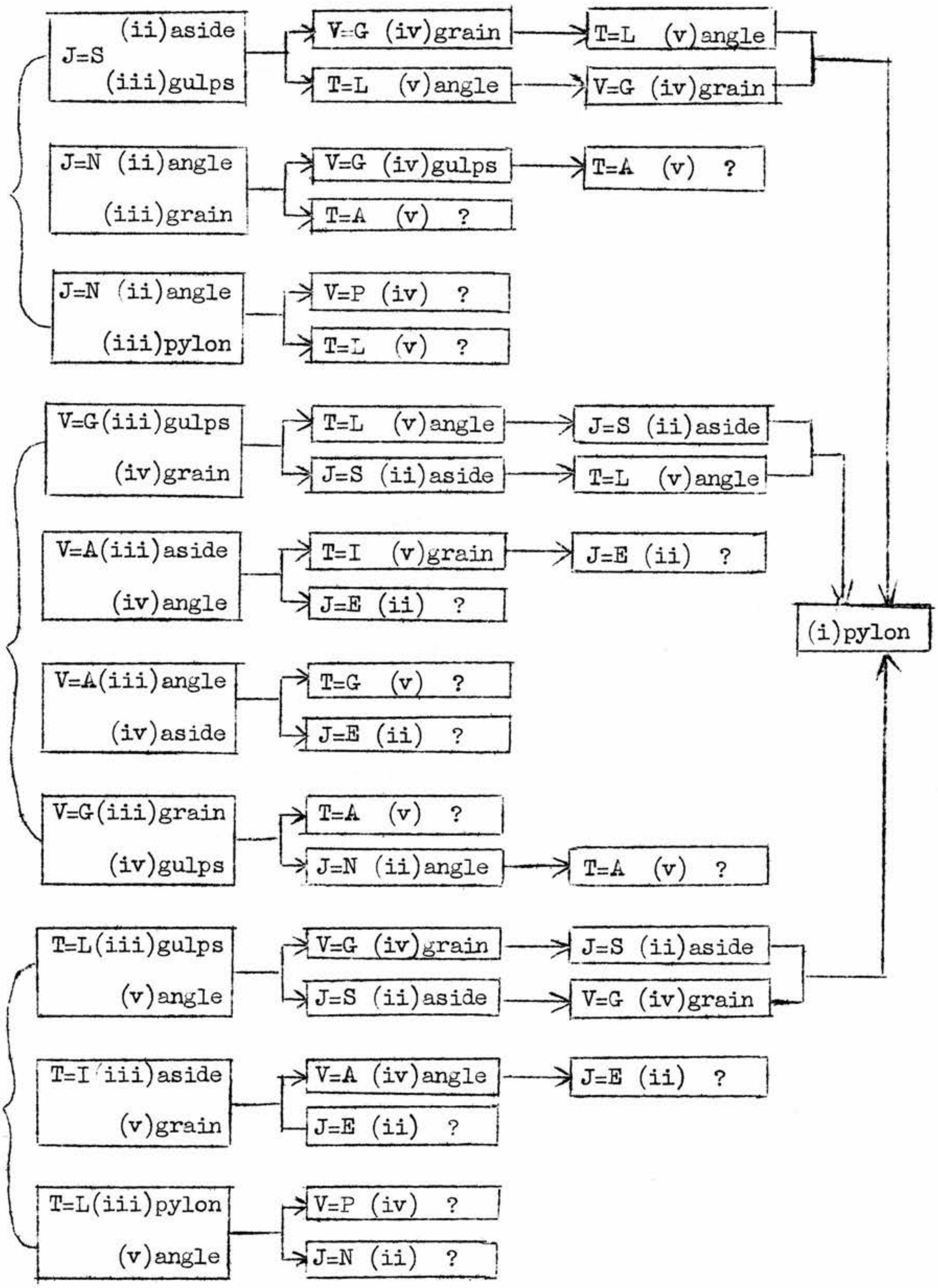
The items in Test II again have three letter pairs, just sufficient for solution, but this time with multiple possible identifications for each. Whichever letter pair one begins with, there is the possibility of starting out on a wrong path, e.g.

```
(1)GRAIN   (2)ASIDE   (3)PYLON   (4)GULPS   (5)ANGLE

                  - - - - -    ..............(   )
                  - J - - -    ..............(   )
                  V - T - J    ..............(   )
                  V - - - -    ..............(   )
                  - - - T -    ..............(   )
```

<u>Item II</u>

It may be noted that the pattern of coded letters is exactly the same as for Item I. This was arranged so that the perceptual impact of the items as stimulus displays might be the same for both sorts of item with just sufficient data. The only variable on which they differ is in presence versus absence of false leads, this having come about by the substitution of the word ANGLE, in Item II, for the word EAGLE in Item I. The various routes to solution and false leads for Item II are as follows:



The over-all probability of starting out on a wrong path - calculated in the manner outlined on p. 81 - is equal to 0.69.

In Test III the same sets of words are used as in Test II, but the items have three extra (redundant) letter pairs and all of them open up additional false leads, e.g.

(1)GRAIN	(2)ASIDE	(3)PYLON	(4)GULPS	(5)ANGLE	
-	-	-	-	H	..............( )
-	J	Z	-	-	..............( )
V	-	T	-	J	..............( )
V	-	M	Z	H	..............( )
M	-	-	T	-	..............( )

Item III

The possible starting points for the extra letter pairs (the H's, Z's and M's) are as follows:

H's	Z's	M's
H=N (i)pylon (iv)grain *	Z=I (ii)aside (iv)grain *	M=A (iv)grain (v)aside *
H=N (i)grain (iv)pylon	Z=L (ii)pylon (iv)angle	M=A (iv)grain (v)angle
H=E (i)angle (iv)aside	Z=L (ii)gulps (iv)angle	M=G (iv)angle (v)gulps
H=E (i)aside (iv)angle		

Of these starting points, all except the three with asterisks are false leads. For Item III the over-all probability of starting out on a wrong path is 0.71, i.e. the presence of redundant letter pairs has not here led to a decrease in the probability of starting out on a wrong path (it has in fact led to a slight increase). For all the Test III (redundant) items the probability of starting out on a wrong path is of the same order as for the corresponding nonredundant items of Test II (although usually a little higher). The actual probability values are shown below, along with those for Test I.

Item	Test I Just sufficient data	Test II Just sufficient data	Test III Extra redundant letter pairs
1		0.56	0.61
2		0.61	0.64
3		0.58	0.67
4		0.67	0.67
5	P = 0.00	0.66	0.73
6	in every case	0.69	0.71
7		0.69	0.71
8		0.67	0.68
9		0.72	0.73
10		0.72	0.74
11		0.74	0.74
12		0.63	0.74

(Note: The probability of starting out on a wrong path for the items of Test I is zero only if the subjects begin with the triangled letter pairs as instructed. But even if they do not, there are comparatively few false leads - nowhere near as many as for Tests II and III).

In line with the hypothesis that redundancy either aids or hinders solution according as it leads to a decrease or to an increase in the probability of starting out on a wrong path, predictions were made as follows:

(i) Test III will be of the same general level of difficulty as Test II, and certainly not any less difficult*

* Although it is not recommended research strategy to predict a null result, it does seem in order here. Redundant items have invariably been found less difficult than nonredundant up to the present: to show them not to be less difficult here would argue persuasively in support of the stated hypothesis.

(ii) Test III will be more difficult than Test I

It was also predicted, owing to the presence of false leads, that

(iii) Test II will be more difficult than Test I.

The tests were administered one to each of three tutorial groups. The groups were assumed not to differ with respect to variables relevant to the solving of code items, but no check was made on this. Each test was prefaced by two sample items analogous to the items in the test proper. 20 minutes working time was allowed. The results are set out in Table 5, along with correlation data to be discussed at a later stage (see p. 113 ff.). Copies of the tests are included in Appendix I.

	Test I Just suffic. data and no false leads (N=20)	Test II Just suffic. data, with false leads (N=17)	Test III Redundant data, with false leads as for Test II (N=27)
Number whole items correct (max. score=12)	$\bar{X}$ = 11.8 R = (11-12)	$\bar{X}$ = 7.2 R = (0-12)	$\bar{X}$ = 6.0 R = (1-11)
Number single words attempted (max. score=60)	$\bar{X}$ = 60.0 R = ___	$\bar{X}$ = 42.2 R = (13-60)	$\bar{X}$ = 37.6 R = (10-60)
Number single words correct (max. score=60)	$\bar{X}$ = 59.6 R = (58-60)	$\bar{X}$ = 37.7 R = (2-60)	$\bar{X}$ = 32.7 R = (8-58)
Number errors (words incorrect)	$\bar{X}$ = 0.2 R = (0-2)	$\bar{X}$ = 4.4 R = (0-27)	$\bar{X}$ = 4.8 R = (0-22)
Error ratios	0.7%	10.5%	12.8%
Correlation with Group Test 33	Not computed due to lack of scatter on codes test	r = 0.23 (N=14)	r = 0.43 (N=24)
Correlation with embedded figures test		r = 0.20	r = -0.01

Table 5

The over-all difference between groups is significant for all measures (Median tests extended for k independent samples; all P values < 0.001, 1-tailed). Mann-Whitney U tests then reveal that all differences between Tests I and II and Tests I and III are significant (all P values < 0.001, 1-tailed). None of the differences between Tests II and III is significant. Hence predictions (i)-(iii) have all been confirmed. Test I proved ridiculously easy. All subjects finished within 10 minutes, and testing was then curtailed. There was virtually no scatter of scores, so that the summary statistics for Test I are somewhat meaningless. What the test has demonstrated is that items with just sufficient data for solution are not difficult *per se*, which is of interest in relation to the speculations of Wohlwill (1962, p. 99) as to whether the average adult is capable consistently of operating at the level where redundancy is reduced to an absolute minimum - typically zero. Although the above results are for subjects of above average ability they suggest that, for code items at least, the average adult would have no trouble operating at the "rarefied level" of zero redundancy *provided* there are no false leads present. When false leads are introduced - as in Test II - difficulty is experienced by undergraduate subjects, and would presumably be even moreso for the "average adult". The Study 4 results bring it out once again that sheer number of stimulus elements is not all-important as a determinant of task difficulty. The crucial thing in this case is whether the stimulus elements - be they many or few - happen to open up false leads. These problem items conform very nicely, in fact, to the notion espoused in particular, of late, by Newell, Shaw and Simon (e.g. 1958, 1962) that the problem solving process consists of a search for a solution in a "space" of possible solutions, and the smaller the set of possible solutions that has to be searched the easier is the problem and the more rapidly will it be solved. This notion of problem solving as a search in a space of possible solutionsis taken up in the studies involving individual testing (Studies 5-9).

The study 4 results have confirmed that redundancy is helpful to the extent that it leads to a decrease in the over-all probability of starting out on a wrong path. Is this however its only role ? Scrutiny of the items used reveals that redundancy might also be helpful due to its allowing the definite identification of pairs of coded words independently of the other coded words in the item - something not possible for corresponding nonredundant versions. Take for example the following:

(1)FROWN (2)HARPS (3)SHOPS (4)SPEAR (5)WRATH

Just sufficient data	Three extra letter pairs
- X - - -	U X - Q I
- B X - -	- B X - U
- - - - -	I - - - -
- B G - -	- B G - -
- - G - -	- - G Q -

For the version with just sufficient data either:

(i) X = R, with the first coded word FROWN and the second HARPS, or
(ii) X = R, with the first coded word WRATH and the second HARPS, or
(iii) X = A, with the first coded word HARPS and the second WRATH

Which of these possibilities is correct cannot be decided without taking further coded words into account. For the version with three extra (redundant) letter pairs the first two coded words <u>can</u> be identified without taking any other coded word into account, due to these two coded words sharing not only a pair of X's but also a pair of U's. The U's are in first and last position respectively, and rule out alternatives (i) and (ii). Hence X = A and the first coded word is HARPS and the second WRATH. The role of redundancy here is not so much to decrease the probability of starting out on a wrong path as to enable wrong paths more quickly to be identified and abandoned. There is however also the possibility of pairs of coded words that share more than one coded letter pair slowing down, rather than facilitating, solution. Suppose one begins the item above with the B's. Two possibilities arise:

(i) B = R, with the second coded word FROWN and the fourth WRATH
(ii) B = R, with the second coded word WRATH and the fourth FROWN

If (i) is correct then X = O and the first coded word, since it has an X in second position, must represent a word in which the second letter is O. No such word is listed. Hence (ii), not (i), must be correct. If (ii) is correct then X = A and the first coded word will be HARPS. This being so, U = H. One may then seek to make further headway by locating the other H. The other H is however at the end of the second coded word, and the second coded word has already been identified. Hence the overlapping of the X's and the U's has not here advanced solution: it has merely confirmed what was already known, at the expense of a certain amount of solution time.

In 7 of the 12 redundant items used in Study 4 there were pairs of coded words sharing two coded letter pairs. The extent to which subjects did or did not capitalise on these shared letter pairs could not be directly assessed, but it seems doubtful that they were fully exploited. In the first place a subject has to see that a pair of coded words has two coded letter pairs in common, and this is not something that is immediately perceptually apparent. In the second place the subject has to employ the strategy of taking two letter pairs into account simultaneously, and this is a strategy that few subjects use. To deal with only one pair of coded letters at a time may seem a pedestrian and low-level way of proceeding, yet it is the way in which the majority even of University students do in fact proceed. The flow charts embodied in this thesis are not artificial in their portrayal of the problem solving process as a succession of separate but interrelated steps. Only with a few of the ablest subjects were there attempts to take more than one letter pair into account simultaneously, and even these subjects usually gave up the attempt when the letter pairs had multiple initially-plausible identifications. Bruner and his colleagues (1956) talk about "cognitive strain" - a taxing of one's powers of memory and inference. The taking of two or more letter pairs into account simultaneously when each has a number of initially-plausible identifications could involve considerable strain on immediate memory, and this is no doubt one reason why the procedure is avoided.* The business of taking a number of things into account

---

* This avoidance occurs in spite of the fact that pencil and paper are provided, and the margins of the test sheets can be used - and are extensively used - as scribbling paper. Had this not been allowed the cognitive strain may well have been too great for subjects other than the very ablest to have solved the items, even taking only one letter pair into account at a time.

simultaneously and organising them together for response is probably also a function of intelligence (as distinct from immediate memory). Piaget has shown that the ability to take various facets of a situation into account simultaneously, rather than centring on one alone, develops with intelligence; and according to Burt (1949, p. 193) the distinctive thing about the intelligent child is "his power to re-arrange a set of related data, and to combine them into a coherent relational whole". Intelligence, for Burt, is essentially a synthesising and organising capacity.

A questionnaire was distributed at the completion of the testing, with questions as below:

(1) In solving the items did you consider all the given letters or only some ?

(2) Did the number of letters you considered vary from item to item ?

(3) Did you check your answers in any way for correctness ? If so, how ?

(4) Did you evolve any methods or rules for solving the items ? (If so, describe)

(5) Did you have any preferred starting point for solving the items (e.g. to look first at words with coded letters in the middle positions; to start with the first coded word and work downwards; to look for words with the same coded letter in first and last positions; etc.) ?

(6) Did you find it helpful - or necessary - to write the various solution words down in addition to filling in their numbers ?

(7) Do you think your skill in solving the items developed with practice ?

(8) Any other comments on the items ?

The replies to these questions shed further light on the role of redundant data in the solving of code items. Replies to questions

(1)-(3) were as follows:

Tests	Percentage of subjects stating that			
	Some letters used	All letters used	No. of letters varied	Answers checked
I (N=16)	31.3%	68.7%	0.0%	75.0%
II (N=16)	37.5%	62.5%	0.0%	66.7%
III (N=13)	100.0%	0.0%	84.6%	92.3%

(N's based on number of subjects returning the questionnaire)

That 37.5% of the subjects who did Test II (just sufficient data, with false leads present) claim to have used only some of the letters given would seem due, as in Study 2, to limited vision or a lack of awareness of possibilities, this being in turn due in part to the pressure of time (see pp. 62-3). This may be the case also for the 31.3% of subjects who claim to have used only some of the given letters for the items in Test I (just sufficient data, no false leads). On the other hand, some of these 31.3% of subjects may have happened to have become aware of the fact that some of the Test I items are soluble from two letter pairs alone (see pp. 85-6). No subject who did the redundant items (Test III) claimed to have used all the given letters - in Study 2 12.0% of subjects made such a claim. It was suggested in the discussion of the Study 2 questionnaire results that pressure of time might have been the reason for not all the given letters in the redundant items being used: checking to see if the letter pairs as yet unused are consistent with the obtained solution is wasteful of precious time. This suggestion is quashed, at least as far as the present study is concerned, by the replies to question (3). 92.3% of subjects who did the redundant items report having checked their solutions for correctness, and the majority of them (69.2%) did this by - to quote - "ensuring that the letters I did not use for solving fitted in with the answers I had obtained". (The other 23.1% went no further than to ensure that just one of the letter pairs not used for solving fitted in with the answers). The 12.0% of subjects who claimed to have used all the letters in the redundant items of Study 2 may well also have meant by this that they used less than all for solving, but all in the course of checking their solutions. No enquiry was made in the Study 2 questionnaire about checking of answers for correctness.

In contrast to the 92.3% of subjects who checked their solutions for correctness on the redundant items, only 75.0% of the subjects who did Test I and 66.7% of the subjects who did Test II report having checked their solutionsfor correctness. Checking, for the items with just sufficient data, is a matter of deriving the solution for a second time from the same letter pairs, although they may be taken in a different sequence. Omission of checking or verification - the classic final stage of the problem solving sequence as discussed by Wallas (1926), Johnson (1955), Poincaré and Patrick (cited by Woodworth & Schlosberg, 1954) and, more recently and in a new sort of language, by Newell, Shaw and Simon (1962) - lays one open to the risk of error. That the majority of errors that are made - both structural and executive - are due to a failure to verify, comes out in the studies involving individual testing (Studies 5-9).

The question "Did you evolve any methods or rules for solving the items?" elicited little information of any interest. Taking all 45 subjects who returned the questionnaire as a combined group, 27 of them (60%) said either that they did not develop any methods or rules, or that they followed the method outlined in connection with the sample items. The replies of the other 40% tend to be somewhat vague as to details. 7 people mention locating all the possibilities for every letter pair and then starting with the letter pair with fewest possibilities, while the rest seem to adopt a more "partist" strategy and take one letter pair at a time - try out one possibility for it - and if this proves not to be correct they go on to seek another identification for it until success is eventually achieved. In the words of one subject: "I assumed that one of the possibilities was correct and worked from there, changing the assumption if necessary".

Replies to question (5) - "Did you have any preferred starting point for solving the items ?" - confirmed the existence of preferences, as noted in Study 2, for letter pairs with members in first and last positions or in the same ordinal position. Replies to question (6) - "Did you find it helpful - or necessary - to write the various solution words down in addition to filling in their numbers ?" - were as follows:

Tests	Writing down of the solution words		
	Not necessary	Helpful	Necessary
Test I Just sufficient data - no false leads (N=16)	25.0%	75.0%	0.0%
Test II Just sufficient data - false leads (N=16)	43.8%	37.5%	18.8%
Test III Redundant data - false leads (N=13)	23.0%	23.0%	54.0%

Not unexpectedly, the more false leads there are and the greater the amount of data present in an item the more do subjects find it necessary to write down the various solution words in the course of working out the solution

Replies to question (7) - "Do you think your skill in solving the items developed with practice ?" - were thus:

Tests	Improvement	No improvement
Test I (N = 16)	75.0%	25.0%
Test II (N = 16)	62.5%	37.5%
Test III (N = 13)	53.8%	46.2%

Improvement with practice had been expected, given adequately motivated subjects, in view of the initial unfamiliarity of the item type. In face of an unfamiliar task it is generally found that an appropriate method of attack is developed only after the first few trials (e.g. Chen, 1937). That a smaller percentage of those who did the versions in which false leads are present (Tests II and III) considered their skill in solving the items to have developed with practice than for those who did the test with no false leads may be a reflection of the fact that success on the items in which there are false leads is to some extent dependent on happening not to start out on a false lead. The business of happening or not happening to start out on a false lead has been discussed elsewhere (Campbell, 1963, 1964) under the rubric of "serendipity", i.e. the knack of continually and apparently fortuitously stumbling upon things of interest, value, desirability or profit. Suppose one begins an item with a letter pair for which there are four initially-plausible identifications. Three of them will be false leads. The more false leads the subject pursues prior to attaining the solution - or abandoning the attempt - the fewer the items he will be able to attempt within the time limit and hence the less his chance of a high score. How fortunate then the individual who hits on the correct identification straight off; and how much more fortunate the individual who has this success consistently, as would an individual possessed of serendipity. In this connection, John (1957) argues that a "fortuitous factor of initial choice" is involved in problem solving whenever (i) selection of a certain starting point (e.g. "X") serves to enhance performance, but (ii) there is no a priori reason for starting at this point rather than any other. Both (i) and (ii) apply to the code items, and no amount of practice can be expected to improve one's "luck" in happening to pick the correct identification straight off, for letter pairs with multiple initially-plausible identifications.

To return now to the effects of irrelevancy, it was found in Study 2 that irrelevant letters impede solution although not to a statistically significant degree. In Study 1 on the other hand the difference in difficulty between items with and without irrelevant letters <u>was</u> significant. From the survey of the literature pertaining to irrelevant data in problem solving it was concluded (p. 49) that irrelevant data can be expected to hinder solution, and that the extent of their effect will depend, among other things, on the extent to which they are "obviously irrelevant". An attempt was now made to put this hypothesis to experimental test in relation to the code items. To do this, some measure is needed of the extent to which irrelevant data are "obviously irrelevant". It was predicted that irrelevant data will be the more "obviously irrelevant" the more they are qualitatively unlike the relevant data, and hence the more readily discriminable therefrom.

In Studies 1 and 2 the items with irrelevant letters were of the type below:

(1)ROUGH (2)SLIDE (3)HUMID (4)MELTS (5)LATER

-	F	-	J	-	..............( )
-	Z	J	-	Y	..............( )
C	Q	-	-	-	..............( )
Y	B	-	-	C	..............( )
-	X	-	-	-	..............( )

(1)DELVE (2)KETCH (3)CATER (4)PLUME (5)BRAWN

```
N - - F Z   ..............(   )
U - - Y -   ..............(   )
G - P Q J   ..............(   )
E J Z X -   ..............(   )
H - P D -   ..............(   )
```

How "obviously irrelevant" are the irrelevant letters here ? They are qualitatively similar to the letters relevant to solution in that all are letters of the alphabet, yet there is a ready basis for discriminating between the two: the relevant letters occur in pairs while the irrelevant letters do not, and the irrelevant letters are either all in the same ordinal position or in the same two ordinal positions while the relevant letters fall into no such pattern. The subjects however, to judge by their questionnaire comments, did not appreciate these facts, and it would seem that the irrelevant letters in items such as the above are not very "obviously irrelevant". For the present study items were constructed on the pattern of those below:

(1)PLANS (2)TRIAL (3)AVERS (4)SPITS (5)INERT

```
- - H - O   ..............(   )
- - K - -   ..............(   )
H - - - -   ..............(   )
K - - - O   ..............(   )
- - - - -   ..............(   )
```

Item I

(Just sufficient data)

(1)CREST (2)PRISM (3)ERECT (4)DRIFT (5)IRATE

```
- J H - O   ..............(   )
- - K - -   ..............(   )
H J - - -   ..............(   )
K J - - O   ..............(   )
- J - - -   ..............(   )
```

Item II

(Extra irrelevant letters, "obviously irrelevant")

(1)PRAYS (2)SNIDE (3)ABETS (4)SPIES (5)ICILY

-	-	H	-	O	....	( )
-	M	K	-	-	....	( )
H	-	-	G	-	....	( )
K	F	-	-	O	....	( )
-	-	-	N	-	....	( )

Item III

(Extra irrelevant letters, less "obviously irrelevant" than for Item II)

Solution of each of the Items I- III **rests** on the H's, K's and O's, and the number of initially-plausible identifications for each of these letter pairs is the same for all three items. Items II and III differ from Item I in that each has four extra letters, irrelevant to solution, in addition to the H's, K's and O's. In Item II the J's are irrelevant. Knowing that J = R does not advance solution in any way, since all five uncoded words have an R in second position. In Item III the M, F, G and N are irrelevant: knowing what any one of them stands for advances solution not at all. In line with the hypothesis that irrelevant data are the more "obviously irrelevant" the more they are qualitatively unlike the relevant data and hence the more readily discriminable therefrom, it was predicted that the irrelevant letters in Item II are more "obviously irrelevant" than are those in Item III (they are all the same rather than all being different; and they are all in the same ordinal position instead of being distributed between two different ordinal positions). This being so, it was then predicted that items patterned after Item II will be less difficult than items patterned after Item III, and the difference in difficulty between items patterned after Item II and corresponding items with just sufficient data will be less than the difference in difficulty between items patterned after Item III and corresponding items with just sufficient data. The items with just sufficient data will be the less difficult in each case.

Three sets of 4 items were constructed, with the items in each set patterned respectively after Items I, II and III above. Care was taken to ensure that no letter in Sets II and III deemed irrelevant to solution was negatively or indirectly relevant.* A group of 24 subjects did the items of Sets I and II. Half the group did Set I first, the other half Set II. 8 minutes working time was allowed per set. A second group of 19 subjects did the items of Sets I and III. Half did Set I first, the other half Set III. 8 minutes working time was again allowed per set. The results are set out in Table 6. Also included in Table 6 are correlation data to be discussed at a later stage (see p.113 ff.). Copies of the three sets of items used are included in Appendix I.

---

* Some of the items of Studies 1 and 2 were marred in this way, although it was probably not a serious shortcoming in view of the indication from Study 3 that subjects are reluctant - and some even possibly unable - to use letters as negative instances.

	Group 1 (N=24)		Group 2 (N=19)	
	Set I (just suffic. data)	Set II (extra irrel. letters, all the same)	Set I (just suffic. data)	Set III (extra irrel. letters, all different)
Number whole items correct (max. score=4)	$\overline{X}$= 3.1 R=(1-4)	$\overline{X}$= 2.6 R=(1-4)	$\overline{X}$= 2.8 R=(1-4)	$\overline{X}$= 2.3 R=(0-4)
Number single words attempted (max. score=20)	$\overline{X}$=18.7 R=(10-20)	$\overline{X}$=16.1 R=(10-20)	$\overline{X}$=16.5 R=(8-20)	$\overline{X}$=14.5 R=(7-20)
Number single words correct (max. score=20)	$\overline{X}$=16.5 R=(8-20)	$\overline{X}$=13.9 R=(5-20)	$\overline{X}$=15.3 R=(7-20)	$\overline{X}$=12.9 R=(4-20)
Number errors (words incorrect)	$\overline{X}$= 2.1 R=(0-9)	$\overline{X}$= 2.3 R=(0-10)	$\overline{X}$= 1.2 R=(0-8)	$\overline{X}$= 1.6 R=(0-10)
Error ratios	11.4%	14.0%	7.3%	10.9%
Correln. with Group Test 33	r=0.08 (N=19)	r=0.04 (N=19)	r=0.17 (N=13)	r=0.36 (N=13)
Correln. with embedded figs. test	r=0.01	r=0.34	r=0.02	r=0.44

<u>Table 6</u>

The difference between Sets I and II is significant for number of words attempted and number of words correct (Wilcoxon matched-pairs signed-ranks tests; $P < 0.005$ and $0.025 > P > 0.01$, 1-tailed, respectively). The difference between Sets I and III is significant only for number of words correct (Wilcoxon matched-pairs signed-ranks test; P = 0.025, 1-tailed). As predicted, it is the items with just sufficient data that are the less difficult in each case. The prediction that the difference in difficulty between items of Sets I and II will be less than that between the items of Sets I and III was not confirmed. If anything, the opposite is the case. No direct comparison was made of items with irrelevant letters all the same (Set II) and items with irrelevant letters all different (Set III), but the results suggest that the irrelevant letters in the items with irrelevant letters all the same and the irrelevant letters in the items with irrelevant letters all different both hinder solution of the code items to much the same degree, although not necessarily for the same reasons. If the extent to which irrelevant data impede solution depends on how "obviously irrelevant" they are, then it would seem that irrelevant letters all the same and irrelevant letters all different do not differ in "obviousness" as thought: either irrelevant letters that are all the same are not as "obviously irrelevant" as had been thought, or irrelevant letters that are all different are more "obviously irrelevant" than had been thought. Perhaps irrelevant letters that are all the same are "perceptually obvious" but not "obviously irrelevant". They could be considered perceptually obvious due to their forming an inter-related and cohesive group - a homogeneous vertical column - and Garner (1962, p.340) suggests that people expect in general that variables in the environment that are related to one another will be relevant to the situation at hand.

The hypothesis that irrelevant data will be the more disturbing in their effect the less "obviously irrelevant" they are was one of four hypotheses arrived at from a survey of the literature pertaining to irrelevant data in problem solving. The other three hypotheses were:

(i) that irrelevant data will hinder solution more the more numerous these irrelevant data are

(ii) that irrelevant data will hinder solution more the more they suggest alternative but incorrect solutions

(iii) that irrelevant data will hinder solution more the less practised the subjects are at the task.

Hypotheses (i) and (ii) are not dealt with in any of the code item studies, although (ii) is touched on in a pilot study, using number series items, undertaken during the search for a suitable item type and reported in Appendix II. As for hypothesis (iii), the present study has some bearing on it. In comparing the items of Sets I and II and Sets I and III a counterbalanced design was used, to counteract practice effects. If the groups are now split according to the set of items done first, performance on the items with extra irrelevant letters is found to be consistently worse for the subjects who did these items first than for those who did them second, viz.

	Set II items (N = 24)		Set III items (N = 19)	
	Order done		Order done	
	First (N=12)	Second (N=12)	First (N=9)	Second (N=10)
Number whole items correct	$\bar{X}$= 2.4 R=(1-4)	$\bar{X}$= 2.8 R=(1-4)	$\bar{X}$= 1.4 R=(0-3)	$\bar{X}$= 3.0 R=(1-4)
Number single words attempted	$\bar{X}$=15.9 R=(10-20)	$\bar{X}$=16.8 R=(10-20)	$\bar{X}$=12.2 R=(7-18)	$\bar{X}$=16.5 R=(8-20)
Number single words correct	$\bar{X}$=12.8 R=(5-20)	$\bar{X}$=14.7 R=(6-20)	$\bar{X}$= 9.6 R=(4-18)	$\bar{X}$=15.8 R=(6-20)
Number errors	$\bar{X}$= 3.0 R=(0-10)	$\bar{X}$= 1.6 R=(0-6)	$\bar{X}$= 2.6 R=(0-10)	$\bar{X}$= 0.7 R=(0-4)

The score differences between subjects who did the items with extra irrelevant letters first, as opposed to doing them second, are all significant beyond $P = 0.05$ (Mann-Whitney U tests, 2-tailed), except for number of single words attempted and number of errors on the Set II items. While these results do not show - as hypothesis (iii) would predict - that performance improves with practice on items in which irrelevant data are present, they do show that practice on the general item type, prior to doing the items with extra irrelevant letters, facilitates subsequent performance on these items. Once one has become familiar with the general characteristics of the item type one can then better cope with variations thereon.

A questionnaire was distributed at the completion of the testing, with the questions being much the same as for the questionnaire used in the follow-up study on redundancy (see p. 95)

- except for question (3). The questions were as follows:

(1) In solving the items did you consider all the given letters or only some ?

(2) Did the number of letters you considered vary
(a) from item to item
(b) from the first set of items to the second ?

(3) Did you find any aspect(s) of the items distracting or misleading ?

(4) Did you develop any methods or rules for solving the items ? (If so, describe).

(5) Did you have any preferred starting point for solving the items (e.g. to start with the first coded word and work downwards; to look first at words with coded letters in the middle positions, or in first and last positions; etc.) ?

(6) Did you find it helpful - or necessary - to write the various solution words down in addition to filling in their numbers ?

(7) Do you think your skill in solving the items developed with practice ?

(8) Any other comments on the items ?

The replies to these questions shed further light on the role of irrelevant letters in the solving of code items. Replies to question (1) were as follows:

	Percentage of subjects stating	
	Some letters used	All letters used
Set II items - extra irrelevant letters all the same (N=15)	26.7%	73.3%
Set III items - extra irrelevant letters all different (N=14)	85.7%	14.3%

(N's based on number of subjects returning the questionnaire)

The majority of subjects doing the items with irrelevant letters all the same claim to have used all the given letters in solving the items. This means that they are claiming to have used the letters that are irrelevant, and cannot advance solution in any way. For the items with irrelevant letters all different the majority of subjects state that they used only some of the given letters. Replies to question (2) reveal that the letters not used were the ones that are all different and irrelevant to solution. Why should so many subjects try to use irrelevant letters that are all the same, but so few seek to use irrelevant letters that are all different ? The only explanation that suggests itself from the questionnaire replies is that subjects assume coded letters to be the more useful the more frequently they occur. They assume that if letters occurring twice are useful, letters occurring more than twice must be even more so. Hence the attempts to use the irrelevant letters that are all the same, since they occur four times within the one item. With irrelevant letters that are all different each occurs only once, i.e. less frequently than the letters of the letter pairs. Hence they are the more readily ignored. If this assumption that coded letters are the more useful the more frequently they occur were the only determinant of item difficulty affecting Set II and Set III items differently, one would expect the Set III items to be considerably less difficult than the items of Set II. There is no indication that this is so. Hence there must be some other determinant(s) of item difficulty also affecting Set II and Set III items differentially. Possibly there are perceptual factors operating. Irrelevant letters that are all different are not so easy to "find" and, while subjects are not so prone to try and use these letters, in order for them to be rejected from consideration they must first of all be found. Irrelevant letters that are all the same are the more readily seen but, once found, subjects *do* try to use them. If this is, in fact, what happens, the conclusion (p. 106) that irrelevant letters that are all the same and irrelevant letters that are all different hinder solution of the code items "to much the same degree, although not necessarily for the same reasons" may be amended to read "*not* for the same reasons".

The question "Did you find any aspect(s) of the items distracting or misleading ?" was posed in order to see whether it would elicit comment on the irrelevant letters, but little such comment was forthcoming. 50% of subjects said that they did not find any aspects of the items distracting or misleading, while the most frequent complaint of the other 50% of subjects was the presence of multiple initially-plausible identifications for various of the letter pairs. Replies to question (4) confirmed that, in general, subjects follow the method outlined in connection with the sample items. Replies to question (5) confirm - yet again - that subjects prefer to begin with letter pairs with members in first and last positions, or in the same ordinal position, even if this means beginning with the coded words at the bottom of the list. In reply to question (6) there was general agreement that it is helpful to write down the various possible letter and word identific- ations as one proceeds, in order to keep track of what one is doing, and 12% of subjects considered this not merely helpful but absolutely essential for solution to be achieved. 63% of subjects thought that their skill in solving the items developed with practice; 27% thought not. All of these 27% were subjects who did the items with irrelevant letters subsequent to doing those without any extra irrelevant letters.

In conjunction with all the code items used in Study 4 - those with just sufficient data, those with extra redundant letters, and those with extra irrelevant letters - an embedded figures test was administered (see Appendix I). It consisted of four pages of items, with each page separately timed. 1, 3, 3 and 4 minutes were allowed respectively per page. The items come from a test based on Gottschaldt's* embedded figures, and used by Thurstone (1944) in his factorial study of perception. In addition, intelligence test scores were available for most of the subjects on Group Test 33. In Study 1, using the AH5 Test of High-Grade Intelligence, it was

---

* Gottschaldt, K. Über den Einfluss der Erfahrung auf die Wahnehmung von Figuren. Psychol. Forsch., 1926, 8, 261-317.

found that the ability to separate out code letters that are relevant to solution from those that are not seems not, for undergraduate students, to be a function of intelligence as conventionally measured. There was some support for the notion that the intelligent person is one who can work from a bare minimum of cues. The results of the present study again indicate that the ability to separate out code letters that are relevant to item solution from those that are not is not, for undergraduate students, a function of intelligence as conventionally measured, but this time there is no support for the notion that the more intelligent person is one who can work from a bare minimum of cues. Spearman rank-order correlations between scores on Group Test 33 (of intelligence) and number of coded words correct, for code items with varying amounts and kinds of data, appear in Table 5 (p. 90) and Table 6 (p. 105). The N's for these correlation coefficients are based on those subjects for whom scores on Group Test 33 happened to be available. Not one of the obtained correlations differs significantly from zero. Hence, for undergraduate subjects, the code items are discriminating between subjects on the basis of something other than intelligence as conventionally measured.*

Subsequent to Study 1 some speculation was indulged in as to what the ability to screen out relevant from irrelevant code

---

* These undergraduate groups are, admittedly, selected for intelligence in the first place, which militates against the obtaining of high correlations with intelligence test scores. Nevertheless, the range of scores on Group Test 33 was quite extensive : 129-193.

letters and the ability to cope successfully without the support of redundant code letters might be related to, if not intelligence. The screening out of what is relevant from what is not involves a process of analysis, and the notion has long been in vogue that people can be characterised according to whether their predominant mode of functioning is, broadly speaking, "analytic" or "synthetic" (e.g. Jaensch, 1930; Goldstein & Scheerer, 1941; Hanfmann, 1941; Hanfmann & Kasanin, 1942; Witkin, 1950; Witkin et al., 1954, 1962; Gardner et al,, 1959, 1962a, 1962b, 1962c; Karp 1963). Facility in screening out relevant from irrelevant code letters may be related to whether or not one functions in a predominantly analytic mode, and it was to investigate this possibility that the embedded figures test was administered in Study 4 along with the various code tests. The Gottschaldt embedded figures have been used not only by Thurstone but also, and far more extensively, by Witkin and his associates in a form adapted to individual testing. They are held to indicate one's standing on the dimension of "field independence" by measuring one's ability to overcome an embedding context. The irrelevant code letters scarcely form a "context" for the relevant letter pairs, and they would seem to be more "distracting" than "embedding" in the terminology of Witkin and his associates.* It might therefore seem inappropriate even

* In an "embedding" context the various parts of the embedded figure are used as parts also of different figures, whereas in a "distracting" context the figure itself is left intact but is surrounded by considerable extraneous material. Put in another way, an embedding context obscures a figure by changing its nature, while a distracting context obscures it without changing its nature (Witkin et al., 1962).

to consider the possibility of a relationship between scores on an embedded figures test and scores on code items in which there are irrelevant letters. There is however a considerable body of empirical evidence (Gardner & Long, 1962b) to suggest that the dimension of field-independence (or differentiation, as Witkin now calls it) is concerned with coping with irrelevant data of all sorts, even distracting items completely lacking in cohesive organisation, and not just with overcoming embedding contexts. In view of this it is not at all inappropriate to consider the possibility of a relationship between scores on an embedded figures test and scores on code items in which there are irrelevant letters. Spearman rank-order correlation coefficients between number of coded words correct and scores on the embedded figures test appear in Table 5 (p. 90) and Table 6 (p. 105). Two of these correlation coefficients - and only two - are significantly different from zero: that for code items with extra irrelevant letters all the same ($r = 0.34$) is significant at $P = 0.05$ (1-tailed), while that for code items with extra irrelevant letters all different ($r = 0.44$) is significant at $0.05 > P > 0.025$ (1-tailed). Hence it may be concluded that a positive relationship exists between field-independence and success on code items in which there are irrelevant letters. This does not of course imply that field-independence ("analytic ability") is sufficient in itself to ensure success on code items in which there are irrelevant letters: the mere act of analysing out the letters that occur in pairs from those that occur four times, or only once, does not tell one which of these sets of letters is relevant to item solution, or even that only one of them is relevant rather than both. A capacity for judgment is here required, with relevance being assessed in relation to the item structure.

Speculation as to possible correlates of the ability to cope successfully without the support of redundant code letters (intelligence as conventionally measured having proven not to correlate therewith) did not lead very far. Cattell has stated, in connection with nonverbal matrices, (personal communication) that he thinks "the tendency to fill in missing data is quite as much a function of personality factors U.I.21 and U.I.26 as it is of intelligence",* and the hypothesis suggested itself that those who "need" redundancy are those of cautious or timid disposition, who lack confidence and like to check and double-check their answers, rather than being those who are less intelligent. Redundant items, where solutions are overdetermined, provide ample scope for checking, but to engage in such checking means that each item takes longer than would otherwise be the case, and one's total score will therefore be lower. The more reckless person - or he who is confident in his own ability - will arrive at solution to a redundant item using, say, only three letter pairs. He will not bother to check that the other three pairs "fit in", and by thus saving time he will probably end up with a higher score.** These ideas have not been put to experimental

---

* U.I.21 is "Exuberance", and is associated with energy and impetuousness. U.I.26 is "Self Sentiment Control" and is associated with steady, self-absorbed, effective thinking (Cattell, 1957).

** But not necessarily, for he may have erred and, by not checking, the error(s) will not be picked up.

test, but it is scarcely to be doubted that personality and temperament influence the amount of data one takes into account when more than the bare minimum of data are given. That personality and temperament can markedly affect performance on so-called "intellectual" tests has been so widely demonstrated - especially by those of a clinical bent - that it does not require emphasis here.

To summarise the major findings of Study 4:

(i) redundant letters were found to facilitate solution of the code items to the extent that they lead to a decrease in the probability of starting out on a wrong path, but if instead they lead to an increase in the probability of starting out on a wrong path then they impede rather than facilitate solution

(ii) irrelevant letters that are all the same and irrelevant letters that are all different were both found to impede solution of the code items, and to much the same extent, although probably not for the same reasons

(iii) the prediction that irrelevant letters that are all the same would be found "obviously irrelevant" was not confirmed

(iv) the results lent some support to the hypothesis that irrelevant letters impede solution the more the less practised the subjects are at the task

(v) the subjects tended to assume that coded letters are the more useful, the more frequently they occur

(vi) the ability to separate out code letters that are relevant to solution from those that are not was, as in Study 2, found <u>not</u> to be a function of intelligence as conventionally measured, and neither was there any support for the notion that the intelligent person is one who can work from a bare minimum of cues

(vii) a positive relationship was found between field-independence and success on code items in which there are letters irrelevant to solution.

****

(b) Studies involving individual testing

For the final studies on the effects of redundant and irrelevant data on the solving of code items a switch was made from group to individual testing. It was hoped in this way to gain a fuller knowledge of the ways in which subjects tackle the various items. Group testing offers only limited opportunity for the gaining of such knowledge. Inferences can be drawn from the solutions attained - right or wrong - as to how they were arrived at, but since different people are liable to have arrived at the same solution in different ways such inferences are inconclusive and not very informative. The most rewarding sources of information as to the procedures employed in the solving of code items have, for the group testing studies, been (i) the replies to the various questionnaires and (ii) informal comments spontaneously offered by the subjects after completion of the testing. Both these sources of information have their drawbacks. In the first place, interpretation of the questionnaire replies - as the discussion of them has shown only too well - is inferential and often very speculative. Secondly, both sources of information rest on retrospection, and memory - even immediate memory - is notoriously fallible. Subjects may reconstruct their thought processes incorrectly, albeit in the best of faith. Then again, many subjects are quite inarticulate when it comes to giving an account of their thinking. Individual testing allows an opportunity actually to observe what a subject is doing in the course of solving an item, although knowledge gained in this way is again not perfect or complete. With some subjects there was very little on the behavioural level <u>to</u> observe when they were engaged in solving an item. The code items are not problems of that convenient sort - used to advantage by Bruner, Goodnow and Austin (1956) and John and Miller (1957) among others - where thought processes are externalised into a series of verbal responses and/or motor actions. The lack of verbal commentary could have been overcome by asking the subjects to "think aloud", but this method too has its drawbacks. Apart from the standard objection that the demand for verbalisation may alter the nature of the thought processes brought to bear on a problem, to have asked the subjects to think aloud would have handicapped and possibly also have embarassed those not adept at verbalisation. As it happened, it was only a minority of subjects who engaged in <u>no</u> overt activity in the course of solving the items. Most subjects made copious jottings on the test sheets, and many "thought aloud" of their own accord. To those subjects who did not "think aloud" a few judicious questions were posed at the completion of the testing in an attempt to elicit details as to their method of attack on the various items. The questioning was decidedly flexible and permissive, and did not follow any fixed pattern.

The switch to individual testing was accompanied by a change in item format. The uncoded and the coded words are now listed adjacent, viz.

C L O U D	- - J - P	______________
C O U N T	- - - P -	______________
C I D E R	- J Z - -	______________
C H A I N	- - - - Z	______________
C A R T S	- - - - -	______________

The order in which the coded words are listed relative to the uncoded is determined by reference to a table of random numbers. Each uncoded word is to be written in against its coded counterpart. This change was made because, in the group testing studies, most subjects wrote down the solution words in any case, prior to inserting their numbers in the appropriate pairs of brackets. By requiring simply the writing in of words, the possibility has been obviated of executive errors being made in the course of translating the words to numbers. Horizontal listing of the uncoded words was abandoned in favour of vertical, because subjects in the group testing studies regularly converted the horizontal list of words to a vertical one before tackling an item - probably because a vertical ordering simplifies the task of visual search for pairs of letters shared by different words.

Apart from changes in item format there has also been a major structural change: each letter pair in any one item now has the same number of initially-plausible identifications, i.e. there is the same probability of starting out on a wrong path no matter with which letter pair one begins. This means too that the over-all (average) probability of starting out on a wrong path is, for any item, equal to the probability of starting out on a wrong path for each of the separate letter pairs in that item. This is an improvement, from the research point of view, on the items used previously, where the over-all probability of starting out on a wrong path was generally at variance with the probabilities for the separate letter pairs (see for example p. 81 ff.). Having an equal number of initially-plausible identifications for all the letter pairs in any item means that all subjects are on an equal footing, no matter with which letter pair they begin. (It is assumed that the various initially-plausible identifications for any letter pair are equiprobable).

For each of the studies involving individual testing the test items are arranged in a booklet, with each item on a separate page. Each item is separately timed, and the subject is told when to turn over and start and is asked to indicate when he has finished. The instructions are to work as quickly as possible, but to take care not to make mistakes. Should a solution be offered that is not correct, the subject is asked to revise it. The time for revision is added to the original solution time. The stopwatch was used as unobtrusively as possible, in order to minimise any element of stress it might introduce. Subjects for the studies involving individual testing were undergraduates in the Faculties of Arts, Science, and Social Sciences of the University of Edinburgh, without restriction as to year of study. In contrast to the group testing studies, all subjects were anonymous. Five studies were undertaken in all. They follow on from the four group testing studies, and are reported below in chronological sequence as Studies 5-9.

Study 5

Study 5 is primarily a replication study, with regard to redundancy and the number of false leads present in an item. Three items were constructed with just sufficient data for solution, and with probabilities of 0.00, 0.50 and 0.75 respectively of starting out on a wrong path. A further three items were constructed with the same probabilities of starting out on a wrong path, but this time with the data redundant. The items are as follows:

Nonredundant items

F	L	A	I	R	M	-	-	-	-
A	B	I	D	E	-	M	Q	-	-
L	I	G	H	T	Q	-	-	-	S
O	T	H	E	R	-	-	-	S	-
D	E	I	G	N	-	-	-	-	-

Item I

G	R	O	A	N	J	-	-	-	-
E	V	I	C	T	-	-	-	-	-
W	E	A	V	E	-	X	-	-	-
A	C	R	I	D	-	-	-	B	J
T	O	U	G	H	-	B	X	-	-

Item II

A	R	A	B	S	G	Z	-	-	-
I	S	L	A	M	-	-	-	-	-
S	P	A	D	E	Z	-	-	-	-
R	A	L	L	Y	-	-	-	O	-
F	I	L	C	H	-	G	O	-	-

Item III

Redundant items

P	L	A	I	D	C	-	H	-	-
A	R	G	U	E	-	C	M	V	Q
L	I	S	T	S	M	-	-	-	K
B	O	X	E	S	-	-	-	K	H
D	O	I	L	Y	Q	-	V	-	-

Item I'

I	R	O	N	Y	X	F	-	-	-
E	V	I	C	T	-	-	-	B	Q
D	E	L	V	E	-	M	F	-	-
S	C	R	A	P	Q	B	-	Z	X
T	O	A	S	T	-	Z	M	-	-

Item II'

A	R	E	N	A	Y	Z	-	C	-
I	S	L	E	T	X	P	C	-	-
S	L	O	T	H	Z	-	-	-	-
R	A	I	L	S	P	X	-	Q	-
M	I	L	E	S	-	Y	Q	-	-

Item III'

(The spacing and setting out of these items is, in the test booklets, as for the sample item on p. 120). Each of the letter pairs in Items I and I' has only one possible and initially-plausible identification. Since there are no false leads, the probability of starting out on a wrong path is equal to zero. Each of the letter pairs in Items II and II' has two initially-plausible identifications, which means that there is a probability of

1/2 = 0.50 of starting out on a wrong path no matter with which letter pair one begins. Each of the letter pairs in Items III and III' has four initially-plausible identifications, and there is therefore a probability of 3/4 = 0.75 of starting out on a wrong path no matter with which letter pair one begins.* It has already been demonstrated (Study 4) that redundant items and nonredundant items that do not differ with respect to the probability of starting out on a wrong path do not differ in difficulty. There is no reason to suppose that this will not hold true also for individual testing.

---

* For the items with four initially-plausible identifications per letter pair, the starting points for any one particular letter pair are not always mutually exclusive. Take for example Item III. Possible identifications for the G's are I, A, S or R - i.e. four different and mutually exclusive possibilities. The same applies also for the Z's. For the O's on the other hand the possibilities are as follows:

O=L (iv)rally (v)filch	O=A (iv)islam (v)spade
O=L (iv)rally (v)islam	O=A (iv)islam (v)arabs

i.e., here the possibilities are not all mutually independent. This sort of situation was avoided whenever possible, but to do so was not at all easy.

It might also be noted that, strictly speaking, the data in Item I, and other such items where there are no false leads, are <u>more</u> than sufficient for solution (i.e. redundant) <u>if</u> the subject works from the two non-overlapping letter pairs (in Item I the M's and S's). Three letter pairs are nonetheless always given in this sort of item, for the sake of uniformity with the other items where false leads are present and where all three letter pairs <u>are</u> needed for the data to be sufficient for solution.

It was therefore predicted that there will be no difference in difficulty (as measured by mean solution times) between Items I and I', Items II and II', and Items III and III'. 12 subjects were tested, each doing all 6 items. The order in which the items were done was arranged according to two independent Latin Squares, viz.

Subjects	Order of doing items					
	1	2	3	4	5	6
1	I	I'	III'	III	II	II'
2	II	III'	I	II'	I'	III
3	I'	II'	III	II	I	III'
4	III	II	I'	III'	II'	I
5	III'	I	II'	I'	III	II
6	II'	III	II	I	III'	I'
7	III'	I	II	III	II'	I'
8	I	II'	III'	I'	II	III
9	II	III	I'	II'	I	III'
10	I'	III'	I	II	III	II'
11	III	I'	II'	I	III'	II
12	II'	II	III	III'	I'	I

4 sample items prefacedthe test (see Appendix I). Solution of the first 2 was demonstrated to the subject, while the other 2 he was required to do for himself. The results were analysed in the manner set out by Edwards (1960, p. 259ff.) for replication with independent Latin Squares. The outcome of the analysis was as follows:*

---

* Latin Square analyses were used throughout Studies 5-7, and in Study 9. They permit control over the row and column variables and remove the effect of these variables from the error term for treatment comparisons. This is a decided advantage and makes for increased precision when, as here, the rows and columns represent sources of variation liable to be of some importance (individual differences and order of doing the items) but whose effects it is desired to control rather than to assess.

Treatment Means (in secs.)

Nonredundant Items			Redundant Items		
I	II	III	I'	II'	III'
45.75	99.33	164.58	60.17	114.00	172.42

Analysis of Variance - Summary Table

Source of variation	df	SS	MS	F
Squares	1	1,691.68	1,691.68	
Between Ss. in same sq.	10	82,652.71	8,265.27	
Treatments	5	163,368.46	32,673.69	11.18
Order	5	14,719.96	2,943.99	
Squares x Order	5	37,483.40	7,496.68	
Squares x Treatments	5	8,288.90	1,657.78	
Error	40	116,881.66	2,922.04	
Total	71	425,068.79		

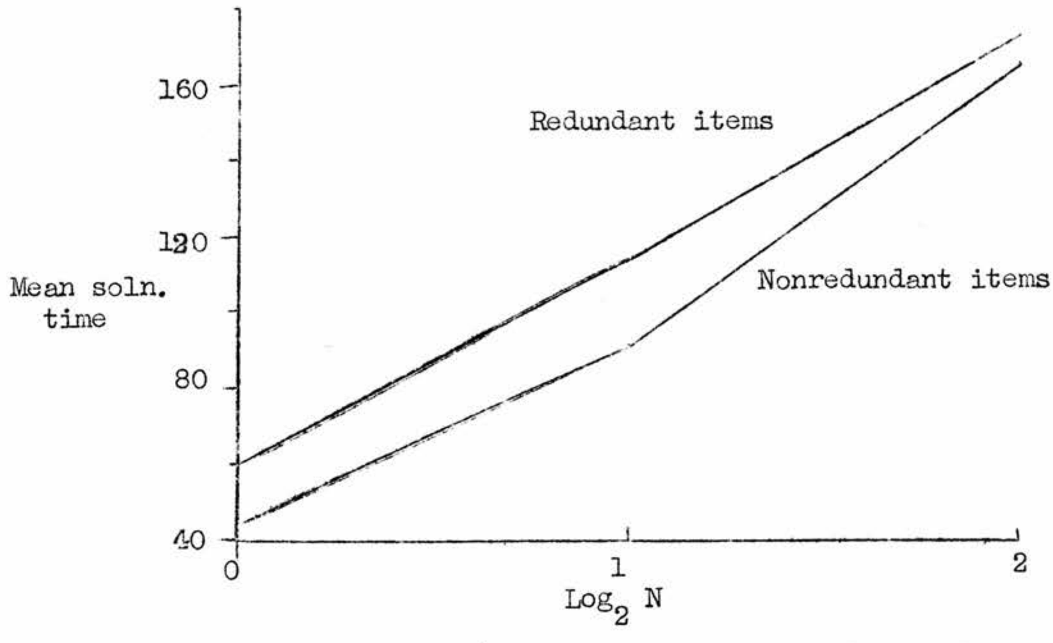
The F for treatments is significant beyond P = 0.01 (df = 5 and 40). This being so, planned orthogonal comparisons were made between the treatment means for Items I vs I', Items II vs II', and Items III vs III' (see Edwards, 1960, p. 140ff.).

Item	Mean ($\overline{X}$)	c_1	c_2	c_3	$c_1\overline{X}$	$c_2\overline{X}$	$c_3\overline{X}$
I	45.75	-1	0	0	-45.75	0	0
II	99.33	0	-1	0	0	-99.33	0
III	164.58	0	0	-1	0	0	-164.58
I'	60.17	1	0	0	60.17	0	0
II'	114.00	0	1	0	0	114.00	0
III'	172.42	0	0	1	0	0	172.42
	Σc	0	0	0	14.42	14.67	7.84
	Σc^2	2	2	2	(d_1)	(d_2)	(d_3)

t_1=0.65 t_2=0.67 t_3=0.36

t_1 and t_2 and t_3 are testing the significance of the difference between means for Items I vs I', II vs II' and III vs III' respectively. None of the t values is significant. Hence the prediction has been confirmed that redundant and nonredundant items that do not differ with respect to the probability of starting out on a wrong path do not differ in difficulty. While none of these differences is significant, the mean solution times for the redundant items are in every case greater than for the corresponding nonredundant versions. This suggests that sheer amount of data (number of letter pairs) may assume importance as a determinant of item difficulty when all other things - and especially the probability of starting out on a wrong path - are held equal (cf. pp. 54 and 60).

Mean solution times for nonredundant and redundant items alike increase with an increase in the probability of starting out on a wrong path. Plotting mean solution time against $\log_2$ number of initially-plausible identifications per letter pair reveals a linear trend for both nonredundant and redundant items:



No. init. plaus. identificns./letter pair:	1	2	4
Log_2 N:	0	1	2

If the horizontal axis were re-labelled

0 1 2

Uncertainty in bits

the graph would look very like those for reaction time studies in relation to stimulus uncertainty, within the framework of information theory (e.g. Hick, 1952; Hyman, 1953; Luce, 1960). Although information theory notions have been rejected in so far as redundancy and irrelevancy (noise) are concerned (see pp. 40-44), the information theory notion of uncertainty (number of available alternatives) <u>is</u> a useful one in connection with the code items. There is every indication that the number of initially-plausible identifications per letter pair is a significant determinant of task difficulty, and its effects are evaluated in Study 8 - the culmination of the research project.

To summarise the findings of Study 5:

(i) redundant and nonredundant items were found not to differ in difficulty when they do not differ with respect to the probability of starting out on a wrong path

(ii) there is a suggestion that sheer amount of data may act as a determinant of task difficulty when all other things are held equal, although its effect is not statistically significant

(iii) there is every indication that the number of initially-plausible identifications per letter pair ~~will be~~ is a significant determinant of task difficulty.

Study 6

Studies 6 and 7 are preparatory to Study 8, the final follow-up study concerned with the effects on item solution of redundancy, irrelevancy, and the probability of starting out on a wrong path. The present study was designed to determine whether irrelevant letters vary in their impact according to the ordinal position in which they occur. That their impact might vary according to the ordinal position in which they occur seemed a possibility, in view of the questionnaire replies (Studies 2 and 4) to the effect that letter pairs with members in first and last ordinal positions are preferred to all others. If there is a tendency to look first of all at letters at the extremes of words, irrelevant letters may impede solution more if they are in first or last positions than if they are in a middling position. On the other hand, irrelevant letters may impede solution more when they are in a middling position, since they are then liable to "break up" the (relevant) letter pairs, e.g. as in Item II below, where the P's (which are irrelevant) intervene between one of the Q's and the other. If irrelevant letters vary in their impact according to the ordinal position in which they occur this will have to be taken into account in making any statement as to the effect of irrelevant data on the solving of code items.

Attention is confined in the studies involving individual testing to irrelevant letters that are all the same, and all in the same ordinal position in any one item. The reasons for this are:

(i) it was planned to use, in the final follow-up study, items in which additional relevant letters and additional irrelevant letters are simultaneously present. Such items are not easily achieved when the irrelevant letters are all different.

(ii) irrelevant letters that are all the same were, in Study 4, found markedly disturbing in their effect. This had not been expected, and it was desired to explore the matter further.

3 items were constructed with just sufficient data for solution and with irrelevant letters in first, third and fifth positions respectively. A further 3 items were constructed, again with irrelevant letters in first, third and fifth position, but this time with extra (redundant) letter pairs as well. The items were as follows:

Nonredundant items

F L O W N	C	J	Y	-	-
F I N E R	C	-	-	-	-
F U N D S	C	-	-	-	H
F R U I T	-	-	-	-	J
F E T I D	C	Y	-	H	-

Item I

W A G O N	Q	-	P	-	-
A N G E R	-	-	-	-	-
L O G I C	-	F	P	-	-
B U G L E	-	-	P	Q	X
B I G O T	F	-	P	X	-

Item II

C L O V E	H	-	-	B	D
A R G U E	-	-	-	-	D
S E R V E	-	-	W	-	-
T I N G E	-	-	-	H	D
S L A T E	-	W	B	-	D

Item III

Redundant items

B O G I E	F	H	-	D	-
B E G O T	F	-	Q	P	Z
B A S I C	F	-	P	Z	J
B L A N K	-	D	-	-	H
B U N K S	F	Q	J	-	-

Item I'

O L I V E	F	J	X	-	-
C R I B S	-	M	-	Z	Q
A B I D E	M	P	X	-	-
T R I A L	Q	-	X	F	P
H O I S T	-	-	X	J	Z

Item II'

B A C K S	T	Y	-	F	G
S H E D S	-	-	J	-	G
H O V I S	Q	M	-	Y	-
P I L E S	J	-	M	T	G
C R O P S	-	Q	F	-	G

Item III'

In every case there is but a single possible identification per letter pair, i.e. no false leads. 6 subjects were tested, each subject doing all 6 items. The order in which the items were done was arranged in accordance with a 6x6 Latin Square. The test was prefaced by 4 sample items (see Appendix I), solution of the first 2 being demonstrated to the subject while the other two he was required to do for himself. The analysis of the results is summarised below.

Treatment Means (in secs.)

Items with irrel. letters first		Items with irrel. letters third		Items with irrel. letters last	
I	I'	II	II'	III	III'
70.0	90.0	82.2	77.0	85.5	89.7

Analysis of Variance - Summary Table

Source of variation	df	SS	MS	F
Treatments	5	1,818.89	363.78	1.08
Order	5	1,883.89	376.78	
Subjects	5	11,092.22	2,218.44	
Error	20	6,709.56	335.48	
Total	35	21,504.56		

(Analysis after Edwards, 1960, p. 255ff.)

The F for treatments is not significant. This being so, it may be concluded that irrelevant letters impede item solution to a like degree whether they are at the beginning, in the middle, or at the end of the words, at least for items having no false leads.* That

* Had the F for treatments been significant it had been planned to use Scheffé's test, for comparison of means for items with irrelevant letters 1st vs. 3rd; 1st vs. 5th; and 3rd vs. 5th.

the F for treatments is not significant further confirms the finding (Studies 4 and 5) that redundant and nonredundant items that do not differ with respect to the probability of starting out on a wrong path do not differ in difficulty.

It having been demonstrated that irrelevant letters impede item solution to a like degree whether they are at the beginning, in the middle, or at the end of words, for items having no false leads, a second lot of testing was undertaken in order to determine whether irrelevant letters have the same impact on item solution, whatever their ordinal position, when false leads are introduced to the items. 3 items were constructed with just sufficient data for solution, and a corresponding 3 items with extra (redundant)letter pairs. For each letter pair in every item there are two initially-plausible identifications, and hence a probability of $1/2 = 0.50$ of starting out on a wrong path. The items were as follows:

Nonredundant items		Redundant items	
B E L O W	C Y - - Q	B L E S T	F - H - Q
B O U N D	C - - - -	B A T O N	F - G M -
B A N D S	C - J - -	B R I C K	F J Q - P
B L U S H	- J - Y -	B U L K Y	- G J P H
B R E A K	C - - Q -	B E N D S	F - - - M
Item I		Item I'	
S A L V O	- X P D -	T U R K S	F J X N -
F I L C H	- - - - -	H A R E M	Q Z - J -
B E L O W	- - P - D	M I R T H	N - X - -
I S L E T	M - P X -	P E R M S	Z G X F -
F A L S E	- M P - -	A P R I L	- Q X G -
Item II		Item II'	
S L U M S	H - - B D	H O V I S	T Y - F G
A V E R S	- - - - D	A R E A S	Q - - J G
M I N U S	- W - - -	I S L E S	J - F Z -
L O V E S	- - W H D	M A U L S	- Q - T G
T A I L S	- - B - D	C H U M S	Y - Z - G
Item III		Item III'	

6 subjects were tested, each subject doing all 6 items. The order in which the items were done was arranged in accordance with a 6x6 Latin Square. The test was prefaced by 4 sample items (see Appendix I), solution of the first 2 being demonstrated to the subject while the other 2 he was required to do for himself. The results were as follows:

Treatment Means (in secs.)

Items with irrel. letters first		Items with irrel. letters third		Items with irrel. letters fifth	
I	I'	II	II'	III	III'
138.0	167.3	127.5	187.2	139.3	126.7

Analysis of Variance - Summary Table

Source of variation	df	SS	MS	F
Treatments	5	17,745.66	3,549.13	0.65
Order	5	52,912.66	10,582.53	
Subjects	5	61,340.33	12,268.07	
Error	20	109,709.35	5,485.47	
Total	35	241,708.00		

The F for treatments is not significant. Hence it may be concluded that for items in which there are false leads - as for items in which there are no false leads - irrelevant letters affect item solution to a like degree whether they are at the beginning, in the middle, or at the end of the words. It may be noted that, once again, redundant and nonredundant items that do not differ with respect to the probability of starting out on a wrong path do not differ in difficulty. It may also be noted that the mean solution times for the items with false leads present, and a probability of 0.50 of starting out on a wrong path, are considerably greater than for the items with no false leads (done by a different, but comparable, group of subjects):

Treatment Means (in secs.)

	Items with irrel. letters first		Items with irrel. letters third		Items with irrel. letters last	
Items with false leads	138.0	167.3	127.5	187.2	139.3	126.7
Items with no false leads	70.0	90.0	82.2	77.0	85.5	89.7

Here is yet a further indication that the number of initially-plausible identifications per letter pair is likely to be found a significant determinant of item difficulty.

To summarise the findings of Study 6:

(i) irrelevant letters that are all the same were found to affect item solution similarly whether they are at the beginning, in the middle, or at the end of the words. This is true both for items in which there are no false leads, and for items in which false leads are present.

(ii) the results provide further (incidental) support for the finding that redundant and nonredundant items that do not differ with respect to the probability of starting out on a wrong path do not differ in difficulty

(iii) there is, as in Study 5, every indication that the number of initially-plausible identifications per letter pair is a significant determinant of task difficulty.

Study 7

Study 7 was designed to determine whether letter pairs with members in the same ordinal position are easier to work with than letter pairs with members in different ordinal positions. Take the following items:

	Item I	Item I'
S M A R T	V - J - -	- - J - -
T R A C K	- - - - D	- - - - D
U N T I E	- - - - -	- V - - -
C L I M B	- - - J -	- - - J -
G R O U P	D - - V -	D V - - -

Suppose one begins with the V's. For Item I two possibilities arise:

(i) V=U and the first coded word is UNTIE and the fifth GROUP

(ii) V=C and the first coded word is CLIMB and the fifth TRACK

For Item I' there are again two possibilities:

(i) V=R and the third coded word is TRACK and the fifth GROUP

(ii) V=R and the third coded word is GROUP and the fifth TRACK

For Item I' the two possibilities merely involve a positional interchange of the same two words. There is no uncertainty as to letter identification: V=R whichever alternative is correct. Also, it would seem an easier task to scan the array of uncoded words for pairs of words with the same letter in second position than for pairs of words with the same letter first and fourth. In view of these various considerations it was predicted that letter pairs whose members are in the same ordinal position will be easier to work with, other things being equal, than letter pairs with members in different ordinal positions.

A set of 4 items was constructed, in accordance with the following scheme:

		Ordinal position of members of letter pairs	
		Same	Different
Probability of wrong path	P=0.50	I	I'
	P=0.75	II	II'

Each item has just sufficient data for solution, as follows:

C L O W N	M	-	U	-	-	C O V E R	X	J	-	-	-
B R A N D	-	-	-	X	-	T E N T S	-	-	-	B	-
S W E P T	-	-	-	-	-	A P R I L	-	-	-	-	-
H A R P S	-	-	U	X	-	E A R T H	J	-	-	-	-
B L E S S	M	-	-	-	-	H O U S E	-	X	-	-	B

Item I | Item I'

F L E S H	J	-	-	D	-	A R A B S	G	Z	-	-	-
P R I S M	-	-	-	-	K	F I L C H	-	-	-	Q	-
F A R C E	-	-	-	-	-	S P A D E	-	-	-	-	-
B R A V E	-	-	-	D	K	I S L A M	Z	-	-	-	-
P E A C H	J	-	-	-	-	R A L L Y	-	G	Q	-	-

Item II | Item II'

16 subjects were tested, each doing all 4 items. The order in which the items were done was arranged in accordance with four independently replicated Latin Squares. The test was prefaced by 4 sample items (see Appendix I), the solution of the first 2 being demonstrated to the subject while he was required to solve the remaining 2 for himself. The results were as follows:

Treatment Means (in secs.)

Items:	I	II	I'	II'
	62.19	162.50	104.31	197.81

Analysis of Variance - Summary Table

Source of variation	df	SS	MS	F
Squares	3	15,423.04	5,141.01	
Between Ss. in same sq.	12	135,885.57	11,323.80	
Treatments	3	174,425.04	58,141.68	17.74
Order	3	3,534.54	1,178.18	
Squares x Order	9	19,060.03	2,117.78	
Squares x Treatments	9	13,639.53	1,515.50	
Error	24	78,677.61	3,278.23	
Total	63	440,645.36		

(Analysis after Edwards, 1960, p. 259ff.)

The F for treatments is significant beyond P = 0.01 (df = 3 and 24). Orthogonal comparisons between treatment means were then made. Not more than three mutually-orthogonal comparisons are possible for four treatment conditions, and the following comparisons were considered the most appropriate:

(i) between items with letter pairs with members in the same ordinal position (Items I and II) as opposed to items with letter pairs with members in different ordinal positions (Items I' and II'), the latter being predicted the more difficult

(ii) between items with a probability of 0.50 of starting out on a wrong path (Items I and I') as opposed to items with a probability of 0.75 of starting out on a wrong path (Items II and II'), the latter being predicted the more difficult

(iii) between Items I and II' as opposed to Items II and I', this testing the significance of the interaction between ordinal position of the members of each letter pair and probability of starting out on a wrong path.

Item	Mean($\bar{X}$)		c_1	c_2	c_3	$c_1\bar{X}$	$c_2\bar{X}$	$c_3\bar{X}$
I	62.19		-1	-1	1	-62.19	-62.19	62.19
II	162.50		-1	1	-1	-162.50	162.50	-162.50
I'	104.31		1	-1	-1	104.31	-104.31	-104.31
II'	197.81		1	1	1	197.81	197.81	197.81
		Σc_2	0	0	0	77.43	193.81	-6.81
		Σc^2	4	4	4	(d_1)	(d_2)	(d_3)
						t_1=2.71	t_2=6.77	t_3=-0.24

t_1 is testing the significance of the difference between mean solution times for items with letters pairs with members in the same as opposed to different ordinal positions. The difference is, as predicted, significant ($0.005 > P > 0.0025$, 1-tailed).

t_2 is testing the significance of the difference between mean solution times for items with a probability of 0.50 of starting out on a wrong path as opposed to items with a probability of 0.75 of starting out on a wrong path. The difference is, as predicted, significant ($P < 0.0025$, 1-tailed).

t_3 is testing the significance of the interaction between ordinal position of members of the letter pairs (same vs different) and probability of starting out on a wrong path. The interaction is not significant.

For the items with letter pairs with members in the same ordinal positions there was a pronounced tendency for the subjects to take two letter pairs into account at a time. This makes for more rapid detection of false leads and for quicker solution. It is a method procedure used but rarely for items other than those in which the members of each letter pair are in the same ordinal position (as in Items I and II on p. 136). For items with letter pairs with members in different ordinal positions the taking of more than one letter pair into account at a time creates too great a cognitive strain for all but the ablest subjects (see p. 94). Since letter pairs with members in the same ordinal position have been found significantly easier to work with than letter pairs with members in different ordinal positions it was considered advisable to discontinue the indiscriminate mixing of the two sorts of letter pair within any one item. For the final study only letter pairs with members in <u>different</u> ordinal positions were used.

To summarise the findings of Study 7:

(i) letter pairs with members in the same ordinal position were found easier to work with than letter pairs with members in different ordinal positions

(ii) the number of initially-plausible identifications per letter pair was found to be a significant determinant of item difficulty (thus confirming the indications from Studies 5 and 6 as to the importance of this variable)

(iii) the interaction between ordinal position of members of the coded letter pairs (same vs different) and probability of starting out on a wrong path was found not to be significant.

Study 8

Study 8 was designed as a full-scale follow-up study, the aim of which was to co-ordinate the findings of the previous studies and to evaluate the effects on item solution, singly and in conjunction, of:

(i) redundant data

(ii) irrelevant data

(iii) probability of starting out on a wrong path.

A 12-item test was constructed, in accordance with the following plan:

Data over and above those just sufficient for solution

	$\overline{R}$		R	
	$\overline{I}$	I	$\overline{I}$	I
P = 0.00	Item 1	Item 2	Item 3	Item 4
P = 0.50	Item 5	Item 6	Item 7	Item 8
P = 0.75	Item 9	Item 10	Item 11	Item 12

$\underline{R}$ = redundant letters (three letter pairs)
$\overline{R}$ = no redundant letters
$\underline{I}$ = irrelevant letters (four letters all the same)
$\overline{I}$ = no irrelevant letters
P = probability of starting out on a wrong path

The items are set out overleaf. 4 sample items were constructed to preface the test (see Appendix I), solution of the first 2 being demonstrated to the subject while he was required to solve the other 2 for himself.

W A G O N	Q	-	-	-	-		
L O G I C	-	-	-	-	-		
B I G O T	-	D	-	-	-		$\bar{R}$ $\bar{I}$
A N G E R	-	-	-	Q	X		
B U G L E	D	-	-	X	-		

Item 1

F R O N D	B	V	-	-	C		
F I N A L	B	-	-	C	-		
F A D E S	-	-	J	-	V		$\bar{R}$ I
F L O U T	B	-	-	J	-		
F I R T H	B	-	-	-	-		

Item 2

B A C K S	T	Y	-	F	-		
P I L E S	-	-	J	-	-		
S H E D S	Q	M	-	Y	-		R $\bar{I}$
H O V I S	J	-	M	T	-		
C R O P S	-	Q	F	-	-		

Item 3

I R A T E	L	W	Q	-	Z		
G R O U P	-	W	-	Q	D		
C R U S T	D	-	N	L	-		R I
P R I C K	N	W	-	Z	J		
F R E S H	-	W	J	-	-		

Item 4

Items with zero probability of starting out on a wrong path, i.e. no false leads.

C R O W D	- - - L -	______	
I R A T E	- - - - -	______	
P R O D S	H - - - G	______	$\bar{R}$ $\bar{I}$
T R U C K	- - - G L	______	
A R I S E	- - - H -	______	

<u>Item 5</u>

F I L C H	- X Q D -	______	
S A L V O	- - Q - -	______	
I S L E T	- - Q - D	______	$\bar{R}$ I
B E L O W	M - - X -	______	
F A L S E	- M Q - -	______	

<u>Item 6</u>

C L O U D	- - J - P	______	
C O U N T	- Q - P G	______	
C I D E R	- J Z G -	______	R $\bar{I}$
C H A I N	- - F - Z	______	
C A R T S	- - Q - F	______	

<u>Item 7</u>

R O U T E	D K L - J	______	
A G A T E	- Q M - J	______	
U N I T E	L D - - J	______	R I
S I N C E	K - Y - -	______	
G R A V E	Y M Q - J	______	

<u>Item 8</u>

Items with a probability of 0.50 of starting out on a **wrong path**, i.e. 2 initially-plausible identifications per letter pair. The different starting points for each letter pair are in every case mutually exclusive (see fn. p. 124). This is true also for the items with a probability of 0.75 of starting out on a wrong path, set out overleaf.

B L E S T	- - I - Q	
B E L O W	- I - - -	
B O U N D	- - - - -	$\overline{R}\ \overline{I}$
B A R D S	- - - C -	
B R O W N	- - - Q C	

<u>Item 9</u>

R A I N S	- - U - -	
S M I T H	- - - Z J	
F R I E D	- - U J -	$\overline{R}$ I
A S I D E	- B U - -	
M O I S T	B - U - Z	

<u>Item 10</u>

R A I L S	J - W - -	
A R E A S	O F X W -	
I S L E S	F - - P -	R $\overline{I}$
S C A N S	- J - X -	
M I N U S	- O P - -	

<u>Item 11</u>

M O U N D	V L K - W	
B O N D S	F - - W P	
N O R S E	K L - P A	R I
C O M E T	- L F - -	
R O B I N	- L V A -	

<u>Item 12</u>

Items with a probability of 0.75 of starting out on a wrong path, i.e. 4 initially-plausible identifications per letter pair.

All the principles of item construction outlined in Section II (pp. 10-27) apply to items 1-12, and a very high degree of control has been achieved over item structure. Each of the items for a given probability value could have been placed in <u>any</u> of the categories $\overline{R}\ \overline{I}$, $\overline{R}$ I, R $\overline{I}$ or R I. Take for example Item 1, in which the coded letters are just sufficient for solution. The solution is as follows:

W A G O N	Q - - - -	l o g i c
L O G I C	- - - - -	b i g o t
B I G O T	- D - - -	w a g o n
A N G E R	- - - Q X	b u g l e
B U G L E	D - - X -	a n g e r

Irrelevant letters can be introduced by giving the coded counterparts of the G's, while redundant letters can be introduced by giving the coded counterparts of the pairs of I's, N's and O's.* If all are given in conjunction, an item with additional irrelevant and additional relevant data will be produced (akin to Item 4). This is the first study in which items with both sorts of additional data - relevant and irrelevant - have been utilised. The position of the irrelevant letters, in the items in which such letters appear in the coded array, varies from item to item, but not in any systematic way. This was considered an asset rather than a drawback since it makes for item variety, and the Study 6 results have indicated that irrelevant letters that are all the same affect item solution similarly whether they are at the beginning or in the middle or at the end of the words.

---

* The O's in LOGIC and BIGOT, or LOGIC and WAGON, but not the O's in BIGOT and WAGON since letter pairs with members in the same ordinal position are no longer permissible in view of the results of Study 7.

8 subjects were tested, each doing all 12 items. Although the items are listed above in serial order (1-12), the order in which they were actually done was decided - for each subject separately - by reference to a table of random numbers. This was to enable the use of a Subjects x Treatments analysis of variance, which allows the assessment of any interaction between any two or all three of the main effects (Redundancy, Irrelevancy and Probability of starting out on a wrong path). To be able to assess the interaction - if any - between redundancy and irrelevancy is of interest since, as pointed out on p. 7, both sorts of data are simultaneously present in everyday problem situations. (Latin Square designs, as used in Studies 5-7, assume interaction effects not to exist). On the basis of the results of previous studies, for both group and individual testing, it was predicted:

(i) that the nonredundant and redundant items will not differ significantly in difficulty since they do not differ, over all, in probability of starting out on a wrong path

(ii) that the items with coded letters irrelevant to solution will be significantly more difficult than the items in which such letters do not appear

(iii) that the probability of starting out on a wrong path will significantly influence item difficulty

(iv) that difficulty will increase as a linear function of $\log_2$ number of initially-plausible identifications per letter pair.

<u>Treatment Means (in secs.)</u>

Nonredundant items	($\bar{R}$)	113.10
Redundant items	(R)	148.00
Non-irrelevant items	($\bar{I}$)	111.58
Irrelevant items	(I)	149.52
Items for P=0.00		62.28
Items for P=0.50		109.50
Items for P=0.75		219.88

Analysis of Variance - Summary Table

Source of variation		df	SS	MS	F
Redundancy	(R)	1	29,225.26	29,225.26	13.48**
Irrelevancy	(I)	1	34,542.09	34,542.09	13.93**
Prob. of false lead	(P)	2	418,645.77	209,322.89	42.70***
Subjects	(s)	7	151,006.99	21,572.43	
R x I		1	5,089.59	5,089.59	2.78
R x P		2	9,245.02	4,622.51	3.78*
I x P		2	15,356.44	7,678.22	1.90
R x s		7	15,173.99	2,167.71	
I x s		7	17,359.16	2,479.88	
P x s		14	68,625.73	4,901.84	
R x I x P		2	3,671.69	1,835.85	1.00
R x I x s		7	12,802.00	1,828.86	
R x P x s		14	17,134.49	1,223.89	
I x P x s		14	56,628.07	4,044.86	
R x I x P x s		14	25,669.45	1,833.53	
Total		95	880,175.74		

* $P \leq 0.05$
** $P \leq 0.01$
*** $P \leq 0.001$

The F for Redundancy is significant beyond P = 0.01 (for df 1 and 7). Hence it may be concluded that the redundant items are significantly more difficult than the nonredundant.

The F for Irrelevancy is significant beyond P = 0.01 (for df 1 and 7). Hence it may be concluded that the items in which there are irrelevant letters are significantly more difficult than those without such letters, which is as predicted.

The F for Probability of starting out on a false lead is significant beyond P = 0.001 (for df 2 and 14), and this would therefore seem to be the most important of the main effects as a determinant of solution time. The critical difference for pairs of treatment means for the variable of Probability of starting out on a wrong path is :

61.65 at P = 0.05 (1-tailed), or

91.86 at P = 0.01 (1-tailed).

The actual differences between pairs of treatment means are

P=0.00 vs P=0.50	47.22
P=0.50 vs P=0.75	110.38
P=0.00 vs P=0.75	157.60

Hence it may be concluded:

(i) that the difference in difficulty - as measured by mean solution time - between items with no false leads as opposed to items with a probability of 0.50 of starting out on a wrong path is not significant

(ii) that the difference in difficulty between items with no false leads and items with a probability of 0.75 of starting out on a wrong path <u>is</u> significant ($P < 0.01$)

(iii) that the difference in difficulty between items with a probability of 0.50 of starting out on a wrong path and items with a probability of 0.75 of starting out on a wrong path is significant ($P < 0.01$).

The prediction having been confirmed that probability of starting out on a wrong path will significantly influence item difficulty, coefficients of orthogonal polynomials were applied (see Edwards, 1960, p. 150ff.) to determine whether, as predicted, difficulty increases as a linear function of number of initially-plausible identifications per letter pair. The prediction is upheld ($P < 0.001$).

One of the interaction F's is (just) significant at P = 0.05, viz. that for redundancy and probability of starting out on a wrong path. Hence redundant data vary in their impact according to the number of false leads present. The means in question are set out below:

	<u>P=0.00</u>	<u>P=0.50</u>	<u>P=C.75</u>
$\bar{R}$	113.38	159.75	405.50
R	135.75	278.25	474.00

The interaction between redundancy and irrelevancy - the interaction which is, as pointed out on p. 145, of greatest interest - is not significant. Hence it may be concluded that redundancy and irrelevancy are independent in their effects. No sign here of redundancy compensating for the effects of irrelevancy, in the way that information theory redundancy compensates for the the effects of noise when both are simultaneously present (see p. 44).

The finding that redundant items are significantly more difficult than corresponding nonredundant items ($P < 0.01$) was contrary to prediction, although not entirely unexpected in view of the Study 5 results (see p. 127). Possibly the fact of irrelevant letters being present, in some items, along with the redundant letters accounts for the difference now being statistically significant, but no definite statement can be made. What <u>can</u> be said is that sheer amount of data (number of coded letter pairs) becomes important as as a determinant of item difficulty <u>when the probability of starting out on a wrong path is held constant</u>. The proviso is important for, as the table above shows, a redundant item in which there are no false leads is, on the average, found less difficult - not more difficult -

than nonredundant items in which false leads are present. Also, a redundant item with a probability of 0.50 of starting out on a wrong path is, on the average, found less difficult than nonredundant items with a higher probability of starting out on a wrong path ($P=0.75$).

To summarise the findings of Study 8:

(i) redundant items were found significantly more difficult than corresponding nonredundant items ($P < 0.01$)

(ii) a significant interaction was found between redundancy and probability of starting out on a wrong path ($P < 0.05$)

(iii) items with coded letters irrelevant to solution were, as predicted, found significantly more difficult than corresponding items in which such letters do not occur ($P < 0.01$)

(iv) probability of starting out on a wrong path was, as predicted, found significantly to influence item difficulty ($P < 0.001$)

(v) item difficulty was, as predicted, found to increase as a linear function of $\log_2$ number of initially-plausible identifications per letter pair ($P < 0.001$)

Study 8 represents, in effect, the culmination of the research project. There is one further study to be reported - Study 9 - but it is somewhat out of the main stream of the research. It is a follow-up to Study 3, and is concerned with the effects of redundant and irrelevant data on the solving of items whose solution demands the use of indirect procedures.

****

Qualitative results

A major reason for switching from group to individual testing was the belief that a fuller understanding could thereby be attained as to the ways in which subjects tackle the general item type, and the different versions thereof. Qualitative information as to the solution procedures used was derived from the following sources:

(i) observation of the subjects' behaviour in the testing situation

(ii) verbal statements made in the course of solving the items by those subjects who spontaneously "thought aloud"

(iii) replies given by those subjects who did not "think aloud" to questions posed by the tester, on completion of the testing, as to how they tackled the items

(iv) general comments elicited from the subjects after completion of the testing.

In conjunction with the switch from group to individual testing the item format was changed so that the uncoded words are listed vertically rather than horizontally. A number of subjects adopted the assumption that a coded word and its uncoded counterpart will not be "on the same level". This assumption is not valid, for the order in which the coded words are listed relative to the uncoded words is decided by reference to a table of random numbers. This brings it about that coded words and their uncoded counterparts <u>do</u> at times appear "on the same level". The adoption by subjects of false assumptions has already been noted (see p. 54). The particular false assumption at issue here is worthy of mention because it can prejudice the subject's chances of a rapid solution. Take the item:

C L O U D	- - J - P	______________
C O U N T	- - - P -	______________
C I D E R	- J Z - -	______________
C H A I N	- - - - Z	______________
C A R T S	- - - - -	______________

Suppose one begins with the P's. Two possibilities arise:

(i) P=N and the first coded word is CHAIN and the second COUNT

(ii) P=T and the first coded word is COUNT and the second CARTS

Since COUNT is second in the list of uncoded words, a subject who assumes that a coded word and its uncoded counterpart will not be on the same level is liable to prejudice his chances of a rapid solution by rejecting (i) at the outset, without giving it a trial. He will only return to it after discovering that (ii) is a false lead.

For all the items used in the studies involving individual testing the different letter pairs in any one item all have the same number of initially-plausible identifications. This a number of subjects found "annoying". They started out with the (quite reasonable) assumption that some letter pairs will be "better bets" as starting points than others. Thus, having found in an item with a probability of 0.75 of starting out on a false lead that one letter pair has no fewer than four initially-plausible identific-ations, they decide to abandon this letter pair in the hope of finding another letter pair with fewer than four initially-plausible identifications.* This strategy does not pay off, since all the

* These subjects might be said to be working in accordance with Zipf's law (Cherry, 1961, p. 100ff). Zipf collected statistics, especially on language, in an attempt to show that many human activities are subject to a minimisation principle - the principle of "least effort".

letter pairs in an item with a probability of 0.75 of starting out on a wrong path have four initially-plausible identifications. Items with a probability of 0.75 of starting out on a wrong path were described as being "very effortful", and many subjects bemoaned the fact that in such items "there are too many possibilities to be dealt with".

Factors contributing to the difficulty of items in which there are many initially-plausible identifications per letter pair include the following:

(i) Strain on immediate memory. Items with many initially-plausible identifications per letter pair "require the retention of certain information while one scouts around among the various possibilities", to quote the words of one subject. Memory is, as has already been noted (p. 36), notoriously fallible. Subjects frequently lost track of which alternatives had already been tried out, and even of which letter pair they had set out to work with.

(ii) Perceptual confusions. For items with many initially-plausible identifications per letter pair, the uncoded words tend to draw on subsets of letters more limited than for items with with fewer initially-plausible identifications per letter pair. Compare these items, from Study 8:

R A I L S	J - W - -	________
A R E A S	O F X W -	________
I S L E S	F - - P -	________
S C A N S	- J - X -	________
M I N U S	- O P - -	________

Frequency count of uncoded letters:

A	C	E	I	L	M	N	R	S	U
4	1	2	3	2	1	2	2	7	1

Probability of starting out on a wrong path = 0.75

B A C K S	T Y - F -	________
P I L E S	- - J - -	________
S H E D S	Q M - Y -	________
H O V I S	J - M T -	________
C R O P S	- Q F - -	________

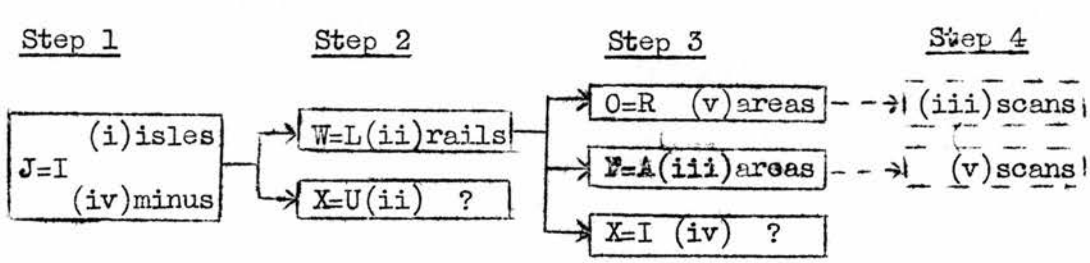
Frequency count of uncoded letters:

A	B	C	D	E	H	I	K	L	O	P	R	S	V
1	1	2	1	2	2	2	1	1	2	2	1	6	1

Probability of starting out on a wrong path = 0.00

For the item with a probability of 0.75 of starting out on a wrong path there are two letters - other than the irrelevant S's - that occur more than twice, viz. the A's and the I's. For items like this subjects report that there are "too many A's" or "too many I's". The confusion is not only perceptual: subjects forget which of the A's or which of the I's formed the letter pair with which they were working. For the item with no false leads such confusion does not arise: no letter - other than the irrelevant S's - occurs more than twice. That the "perceptual look" of an item can be an important determinant of its difficulty is not to be doubted (see for example Elithorn et al., 1960, 1964). A number of subjects commented that for items with a high probability of starting out on a wrong path the various uncoded words are "all too similar".

(iii) <u>Delayed disconfirmation</u>. Suppose one begins the item on p. 152 with the J's, equating J with I. Branch-routes stemming therefrom are as follows:



If one goes from J to X, disconfirmation of J standing for I comes at Step 2. If on the other hand one goes from J to W, disconfirmation is delayed until Step 3 or even later. The dotted lines are intended to indicate that the final coded word has been identified "by default" rather than by looking at the coded letters. If the coded letters are taken into account, disconfirmation of J standing for I comes at Step 4 for the sequences J⟶W⟶O and J⟶W⟶F. If on the other hand the final coded word is identified by default, and verification not then carried out, disconfirmation of J standing for I does not take place at all and an erroneous solution is produced. With items where the letter pairs have many initially-plausible identifications disconfirmation of false leads is frequently delayed. This is in contrast to items having letter pairs with fewer initially-plausible identifications per letter pair, where disconfirmation of false leads is generally immediate, i.e. at Step 2. Delayed disconfirmation of false leads makes the items with a higher probability of starting out on a wrong path the more difficult because time is wasted following false leads beyond Step 2 (and difficulty is being measured by time taken to solve the item), and also because many subjects find it difficult to "backtrack" once they have passed a certain point (usually Step 3). This fits in with such findings as that of Weaver and Madden (cited in Woodworth & Schlosberg, 1954, p. 832) that subjects display great unwillingness to abandon what they have begun and to start all over again once they have reached a certain point.

(iv) Unsystematic attack. Subjects in general take one letter pair at a time and work through the various initially-plausible identifications until the correct one is arrived at. Some subjects however, if they strike a false lead for the letter pair with which they begin, move on to another letter pair. If they strike a false lead for this letter pair too they may move on to yet another, or back to the first. No systematic plan was observable. This sort of procedure can be disastrous when applied to items for which the letter pairs have many initially-plausible identifications: one may strike a false lead for each successive letter pair - and even for each letter pair the second time round - and progress will then be very slow. Memory is important here too, since the subjects who adopted this unsystematic sort of attack generally failed to keep track of the possibilities already tried and the whole procedure became exceedingly disorganised.

(v) Incomplete survey of possibilities. Take again the item:

R A I L S	J - W - -	________
A R E A S	O F X W -	________
I S L E S	F - - P -	________
S C A N S	- J - X -	________
M I N U S	- O P - -	________

Possible starting points for the J's are as follows:

J=A (i) areas
(iv) rails

J=R (i) rails
(iv) areas

J=I (i) isles
(iv) minus

J=S (i) scans
(iv) isles

All but the first are false leads. Time and again did subjects faced with letter pairs having four initially-plausible identifications detect only three of them. This is unimportant if the correct alternative is included among the three, but disastrous if it is not. Suppose the subject decides for the item above that J must be either R or I or S. He tries all three, but none of them "works". He may then go back and scan the list of uncoded words again, but subjects who did this rarely picked up the missing alternative. Many subjects did not even contemplate the possibility of their having missed a possible identification in their initial scanning of the uncoded words. The subject may try the same three possibilities again, on the assumption of having made an executive error on the first run through. Finding again that none of the three identifications works, some subjects then decide the item is insoluble. Assured that it is soluble, they then start again from a different letter pair. Whatever the subjects' reaction, it is clear that an incomplete survey of possibilities hinders solution very markedly if the correct alternative is among the possibilities overlooked. For items with no false leads this never happens, and for items with letter pairs with two initially-plausible identifications per letter pair it is rare, but for items with letter pairs each of which has four initially-plausible identifications it is very common indeed.

Factors (i)-(v) contribute to the difficulty of items for which there are many initially-plausible identifications per letter pair. The introduction of plausible but incorrect alternatives has been used - somewhat intuitively - with various test items as a method of increasing item difficulty. Mazes are an obvious example. The Porteous mazes at the higher age levels have choice points with more false leads and with delayed disconfirmation of the fact that they are false leads, and the "lattice" mazes of Elithorn and his associates (1960, 1964) are similarly constructed.

Reactions to items with just sufficient data for solution as opposed to those with extra data - redundant, irrelevant, or both - were very variable. Items with just sufficient data have one coded word completely blank. Some subjects found this "psychologically offputting" and preferred to work with items with as many coded letters as possible, commenting that "the problems with only a few letters given and a great many spaces are far more difficult than the ones where all the coded letters are given". One subject found it "interesting" - and presumably surprising - that however slight the cues given the words could still be unambiguously decoded. In contrast to the subjects just mentioned, other subjects commented that "the questions with fewest clues are the <u>easiest</u> to work out - confusion results from having too much information". Just as there can be so many initially-plausible alternatives that one does not know where to begin, so too there can be so much data provided that one does not know where to begin. Each item is a perceptual display, and it is well known that beyond a certain point a display load becomes too great for optimum performance. In information theory terminology, channel capacity has been exceeded. Just what it is that distinguishes between subjects who prefer to work with minimal data and those who prefer to work from a great deal of data remains to be seen. Feldman (1964) has suggested that the level of stimulus input that a subject is prepared to tolerate depends on his general level of drive.

Comparing redundant with nonredundant items, a frequent comment on items with six rather than three letter pairs was that the extra data "obscure the structural pattern of the item". It is more difficult to <u>see</u> which letters fall into pairs when there are twelve letters instead of just six. Redundancy also complicates matters by multiplying the number of possible starting points, thereby complicating the process of choosing an appropriate letter pair with which to begin. Subjects commented that nonredundant items seemed easier "because fewer variations are possible in the choice of letters to work with".

For the items with irrelevant letters subjects tended to assume - as in Study 4 - that letters are the more useful the more frequently they occur. Complete failure to appreciate the structural characteristics of the items was not uncommon. One subject given the item

F R O N D	B V - - C	________
F I N A L	B - - C -	________
F A D E S	- - J - V	________
F L O U T	B - - J -	________
F I R T H	B - - - -	________

decided after due deliberation that "B is probably (!) F". Another subject commented that he always looked first at the column with all the letters the same since they at least are "definitely fixed" and the "surest of the lot". (This may reflect the desire for a firm foothold - a definite anchoring point from which to begin). Only 4 of the 8 subjects tested in Study 8 realised that letters such as the B's in the item above are irrelevant to solution. Such letters were variously described as "a pest", "a complicating factor", and "most annoying".

Turning now to some comments that apply to all versions of the items - with redundant data or without, with irrelevant data or without, and with or without false leads - observation of subjects in the individual testing situation confirms that despite response biases, both general (e.g. to begin at the top left hand corner of a stimulus array) and more specific (e.g. to begin with letter pairs with members in first and last ordinal positions), no regular order of scanning of either the coded or the uncoded lists of words exists. In scanning the uncoded words, a few subjects were prone to misreadings - e.g. LOVERS for LOVES, TURK for TURKS, BOUNDS for BOUND. This might or might not interfere with solution, depending on the coded letters given. There was a tendency to assume that coded letters and their uncoded counterparts will be alphabetically "near" to each other. Donaldson (1956) noted a like assumption on the part of schoolchildren, perhaps due to their familiarity with letter series items where nearby letters do often provide the solution. Misreadings and the assumption of alphabetical proximity can be attributed to the code items requiring the subjects to work with stimulus materials (letters and words) conventionally dealt with on the "symbolic" level in terms of their "figural" properties alone.* The letters are to be regarded merely as shapes, and the meaning of the words is irrelevant to item solution. A number of subjects were observed to penalise themselves - without of course intending to do so - by trying at the outset to take all possibilities into account simultaneously. The results is that in seeking the uncoded counterpart of, say, - - J - R they may be seeking a word

---

* Guilford's terms. See for example Guilford et al., 1963.

with an O <u>or</u> an S <u>or</u> an I <u>or</u> a P in third position and an N <u>or</u> a Y <u>or</u> a T <u>or</u> an H in last position. The strain of coping simultaneously with all the overlapping possibilities for all letter pairs proved too great, even for items where the letter pairs have only two initially-plausible identifications, and the attempt to take all possibilities into account simultaneously was soon abandoned.

Errors made on the code items are attributable in the main to a failure to verify one's solution. A solution may be attained that is compatible with a selected subset of letter pairs, and which the subject considers as "satisfactory". It is not in fact satisfactory because it is not the only solution compatible with the selected subset of letter pairs and it is <u>incompatible</u> with the other (unused) letter pair or pairs (see for example p. 154). Errors of this sort can occur for items with only three coded letter pairs (see pp. 16 and 62) as well as for items where the solution is overdetermined. In addition to errors of the sort just discussed an occasional error stems from "loss of hold" (Donaldson, 1963). Given the item

B R E A K	H	J	Z	-	G	______
T O W N S	L	Z	-	-	-	______
E V I C T	-	F	-	H	L	______
M E T E R	G	-	-	-	-	______
C A R V E	-	-	J	F	-	______

one subject proceeded as follows:

<u>Step 1</u>		<u>Step 2</u>		<u>Step 3</u>		<u>Step 4</u>
L=E (ii)evict (iii)carve	→	F=A (v)break	→	J=E (i)mater	→	(iv)towns

Setting J equal to E at Step 3 implies a loss of hold either of the fact of L having already been set as equal to E at Step 1 or of the fact that any given letter has one and only one coded substitute. It may be noted in conclusion that in the course of the studies involving individual testing 48 subjects were tested in all. They solved a total of 304 items.* On only 7 occasions were erroneous solutions offered (which the subject was then asked to revise). This is not to say that false identifications of letter pairs were not made en route to solution, but such errors were picked up at a later stage in all but the 7 cases mentioned. While figures are available for the number of erroneous solutions offered, for the rest of this section on solution methods figures are sadly lacking. This is an unfortunate but unavoidable consequence of the procedures used for collecting the information.

* Study 5:	12 subjects and 6 items per subject	=	72
Study 6:	6 subjects and 6 items per subject plus 6 subjects and 6 items per subject	=	72
Study 7:	16 subjects and 4 items per subject	=	64
Study 8:	8 subjects and 12 items per subject	=	96

Study 9

Study 9 takes up from Study 3, and is concerned with the effects of incorporating redundant and irrelevant data into code items whose solution demands the use of indirect procedures. 3 items were constructed with, as in Study 3, just sufficient data for solution . Each item has only one letter pair (enclosed in triangles), and this letter pair has but a single possible identification. The items are as follows:

I M B U E	-	-	Q	-	-	
S A T I N	-	P	-	-	△F	
B R A V E	-	-	-	-	-	
B E A S T	-	-	-	-	X	
S T A L L	-	J	△F	-	-	

Item I

P U L L S	B	-	-	-	Q	
F I R S T	-	△V	-	Z	-	
C H U T E	-	-	-	N	-	
L O C U S	-	-	-	-	-	
I M A G E	-	-	△V	D	-	

Item II

D E A T H	-	M	-	-	-	
Y E A R S	Z	-	-	-	K	
A N G E R	△O	I	-	-	-	
C H U R N	-	-	-	-	-	
F A I R Y	-	△O	-	-	X	

Item III

The triangled letter pair in each item establishes the identity of two coded words. The identity of the other three coded words can be established only via the use of indirect procedures (apart from guessing). Items I-III provide a baseline for comparison with items having <u>extra</u> letters, relevant or irrelevant.

3 items were constructed comparable in structure to Items I-III, but each with three extra relevant (redundant) letters:

C H U T E	Z	-	F	-	-	______
B A T H E	-	P	-	-	X	______
B L A M E	-	-	-	R	-	______
Y E A S T	-	-	-	-	D	______
B E A T S	Q	J	X	-	-	______

Item I'

R A C K S	B	-	-	-	X	______
C H E S T	-	M	-	V	-	______
G R A T E	L	F	-	Q	-	______
K N E A D	-	-	-	-	-	______
D R O W N	-	I	M	J	-	______

Item II'

B E N C H	-	M	-	-	J	______
T O P A Z	U	-	-	-	F	______
A P P L E	K	Q	-	-	X	______
C H I R P	-	-	D	-	-	______
S A I N T	-	K	-	G	Y	______

Item III'

All the letters in Items I'-III' can be used, at some stage, in arriving at solution, but not all of them need be used, i.e. the solution of each of the items is overdetermined, and the data are redundant.

A third set of 3 items was constructed, again comparable to Items I-III, but each with three extra irrelevant letters. The letters irrelevant to solution have, in the items overleaf, been underlined for ease of reference. Knowing what any one of them stands for does not advance item solution in any way. Each is a "dead end".

F U D G E	H - V - -	
B A T C H	- M - - K	
P L A N T	- - - - -	
L E A S T	- Q - - X	
R O A D S	T J K - -	

Item I''

L U R L Y	G M - - F	
J O I S T	Q Z - V -	
Y O U T H	- - - N -	
L O T U S	- - - - -	
O P E R A	A - Z C -	

Item II''

T E A C H	- M - - -	
Y E A S T	Z - - N B	
A P T L Y	I O - J -	
S H A R P	- - - - -	
D A R K Y	G I - - X	

Item III''

(The spacing and setting out of the items in the test booklets is as for the sample item on p. 120).

In Study 3 it was found that many First Ordinary Psychology students were not able - or not willing to make the effort - to cope with items whose solution demands the use of indirect procedures. If the effects of redundant and irrelevant data on this sort of item are to be assessed, it is important to have subjects who can at least cope with the item type. Final year B.Sc. (Hons.) students were chosen as the most likely candidates (see fn., p. 77). 12 subjects were tested in all, each doing all three sets of items (I-III; I'-III'; and I''-III''). The order in which the sets were done was arranged in accordance with four 3x3 balanced Latin Squares. The test was prefaced by three sample items (see Appendix I), solution of the first two being demonstrated to the subject while the third he was required to do for himself. The results were analysed in the manner set out by Cochran and Cox (1957, p.134ff.). Mean solution times per item (adjusted for residual effects) are as follows:

Treatment Means (in secs.)

Items I-III (just suffic.data)	Items I'-III' (redundant letters)	Items I''-III'' (irrelevant letters)
207.23	331.69	728.90

It may be noted that these figures are very much greater than the mean solution times found in Studies 5-8 for items with varying amounts of data not demanding any use of indirect procedures. The results of the analysis of variance were thus:

Analysis of Variance - Summary Table

Source of variation	df	SS	MS	F
Subjects	11	280,235.99	25,475.00	
Order (w/in squares)	8	312,143.76	39,017.97	
Treatments:				
Direct effects (unadj.)	2	458,197.04		
Residual effects (adj.)	2	6,755.01	3,377.50	0.10
Residual effects (unadj.)	2	108,630.88		
Direct effects (adj.)	2	356,320.36	178,160.18	4.99
Error	12	428,480.25	35,706.69	
Total	35	1,383,935.37		

The F for the direct effects of the treatments (adjusted) is significant ($0.05 > P > 0.01$). The F for the treatment residual effects is not significant. The critical difference for pairs of treatment means (adjusted for residual effects) is:

189.84 - for P = 0.05, 2-tailed

267.89 - for P = 0.01, 2-tailed

Obtained differences are as follows:

Items I'-III'	vs Items I-III	124.46	
Items I''-III''	vs Items I'-III'	397.21	($P < 0.01$)
Items I''-III''	vs Items I III	521.67	($P < 0.01$)

Hence it may be concluded that additional relevant (redundant) letters impede solution of code items demanding the use of indirect procedures, but not to a statistically significant degree, while irrelevant letters significantly impede solution. Both results are in line with those for the "directly" soluble items.

The redundant letters in Items I'-III' act as extra, complicating detail. This is not surprising, for despite their being potentially useful they must be used in the right connection. Take for example Item I'. In addition to the letters occurring in like positions to the letters in Item I there occur also a Z, a Q and an R, viz.

I M B U E	-	-	Q	-	-	
S A T I N	-	P	-	-	F	
B R A V E	-	-	-	-	-	
B E A S T	-	-	-	-	X	
S T A L L	-	J	F	-	-	

Item I

C H U T E	Z	-	F	-	-	
B A T H E	-	P	-	-	X	
B L A M E	-	-	-	R	-	
Y E A S T	-	-	-	-	D	
B E A T S	Q	J	X	-	-	

Item I'

The pair of X's establishes that the second coded word is YEAST and the fifth BATHE. Since Q J X - - represents BATHE, Q stands for B. That Q stands for B is a useful piece of knowledge only in relation to the Z, i.e. the first coded word cannot represent a word beginning with B, which eliminates BEATS and BLAME. That Q stands for B is of no use at all in relation to the D, the R or the F.

The markedly adverse effect caused by the introduction of letters irrelevant to solution can be interpreted in relation to the hypothesis, drawn from the survey of the literature, that irrelevant data will be the more disturbing the less "obviously irrelevant" they are. There is no ready basis in Items I''-III'' for discriminating between letters relevant to solution and those not (unlike the "direct" items where letter frequency is a cue - albeit not one that is widely used). In this particular respect the irrelevant letters in Items I''-III'' are akin to the information theory concept of noise. There is no basic structural difference between signal and noise, the difference being solely in terms of some source of reference that has been agreed on in advance.

Although mean solution times for the items used in Study 9 were considerably higher than for the items used in Studies 5-8, all 12 subjects did eventually solve each item. This is somewhat at variance with the results of Study 3, where many subjects seemed truly unable to employ the sort of procedures appropriate to solution of the indirect code items. The discrepancy in results may be due to the switch from group to individual testing:

subjects in the group testing situation may not have been willing to make the effort required for the solving of the indirect items, whereas a desire not to make a poor showing - and to please the tester - were almost certainly operative in the individual testing situation. Subjects tested individually had the advantage of more sample items, and more detailed explanation of the task. They were, in addition, specially chosen as candidates likely to be adept in the use of procedures based on negative instances. While they proved able to use such procedures, they cannot be said to have been adept in their use. There was a pronounced tendency to convert negative statements to positives, unnecessarily. Take the item:

P U L L S	B - - - Q	________________
F I R S T	- V - Z -	________________
C H U T E	- - - N -	________________
L O C U S	- - - - -	________________
I M A G E	- - V D -	________________

The second coded word is PULLS and the fifth is CHUTE, with V=U and Z=L and D=T. Whatever the letters B, Q and N stand for, they cannot stand for U or L or T. Hence the uncoded counterpart of the first coded word cannot begin with a U, an L or a T. This eliminates LOCUS. Nor can the uncoded counterpart of the first coded word end with a U, an L or a T, which eliminates FIRST. The first coded word must therefore represent IMAGE. Rather than proceeding in this manner most subjects argue that the first coded word does not represent a word beginning with U or L or T and so must represent a word starting with F or I. Nor can the first coded word represent a word ending in U, L or T , which means that it must represent a word ending with S or E. Only IMAGE fits both requirements. In arriving at this conclusion the negative $B \neq$ U or L or T has been transformed to the positive B = F or I; and the negative Q $\neq$ U or L or T has been transformed to the positive Q = S or E. It is as though, as Donaldson (1959) has suggested, negatives are in some way "distrusted".

To summarise the findings of Study 9:

(i) extra relevant (redundant) letters impede solution of code items demanding the use of indirect procedures, but not significantly so

(ii) letters irrelevant to solution significantly impede solution of code items demanding the use of indirect procedures

(iii) all subjects (Final Year Honours B.Sc. students) proved able to cope with indirect procedures based on negative instances, but they did not find the task easy.

V. SUMMARY AND CONCLUSIONS

It is a general feature of psychological studies of human problem solving that the subject is presented with data sufficient for solution, and no more nor any less than these. This is in contrast to problem situations as met with in "real life", where one has to select what are relevant from amid a great welter of data and disregard all the rest. The aim of this thesis has been to determine and account for the effects of incorporating into a problem situation data over and above those logically sufficient for solution. These additional data, as regards their logical status, are either relevant or not relevant to solution. Additional relevant data are such as are able to be used in arriving at solution, but their utilisation is not necessary. They are superfluous and introduce redundancy. Additional irrelevant data have no bearing on solution and are not able to be used in arriving thereat. The problems used in the research were a type of code item adapted from the Moray House Tests of Verbal Reasoning. The early studies employed group testing, and performance was assessed in terms of number of items completed in time t. The final studies employed individual testing, with performance being assessed in terms of time taken to complete each item.

For the items "directly" soluble by the matching up of coded with uncoded pairs of letters redundant data were found, other things being equal,

(i) to facilitate solution to the extent that they lead to a decrease in the over-all probability of starting out on a wrong path

(ii) to impede solution to the extent that they lead to an increase in the over-all probability of starting out on a wrong path

(iii) to impede solution, although not always significantly, if the probability of starting out on a wrong path is the same for redundant as for nonredundant items. Limited support was found for the notion that the intelligent person is one who can work from a bare minimum of cues, i.e. who can do without redundancy.

For items demanding the use of indirect procedures redundant data were found to impede solution, although not to a statistically significant degree. This is in line with the results for the directly soluble items, since the redundant and nonredundant "indirect" items can be regarded as having the same over-all probability of starting out on a wrong path.

Irrelevant data were found significantly to impede solution of direct and indirect code items alike. The effect of letters irrelevant to solution is particularly disturbing for the indirect items, where there is no ready basis for discriminating between letters relevant to solution and those not. The extra irrelevant letters act as complicating detail. There was no evidence that the ability to screen out relevant from irrelevant data is a function of intelligence (as measured by the AH5 Test or the Group Test 33), but a positive relationship was found between field-independence and success on items in which irrelevant data appear. There was a suggestion that irrelevant data are the more disturbing in their effect the less practised the subjects are at the task. Redundant and irrelevant data were found to be independent in their effects.

Probability of starting out on a wrong path was found more significant, as a determinant of item difficulty, than either redundancy or irrelevancy. Other things being equal, the (empirical) difficulty of direct code items increases as a linear function of $\log_2$ number of initially-plausible identifications per letter pair. Each item can be regarded as a maze through which there are a number of paths, possibly overlapping, plus a number of false leads. The task is to discover one of the subset of "correct" paths. A significant interaction was found between probability of starting out on a wrong path and presence versus absence of redundancy, but the form of the relationship remains to be determined.

Errors on the direct items were few, and sprang mainly from a failure to ensure that the proffered solution was compatible not merely with the subset of letters used but also with the letter pair(s) not used in its attainment. Errors on the indirect items were far more prevalent - at least for the group testing situation. The subjects experienced considerable difficulty with the indirect items, and it was suggested that this was due to the demand made by these items for data transformations that involve a change of logical subject, and/or the need to work with equivalence classes that are negatively defined and purely verbal.

The implications of the research for test construction are manifold, e.g. one way of increasing item difficulty is to include in the item redundant data that lead to an increase in the over-all probability of starting out on a wrong path. It is clear that seeming minutiae of item structure can be of importance in determining item solution. Thus the use of letter pairs with members in the same rather than in different ordinal positions leads to a significant decrease in difficulty. The findings of the research can also be applied to educational practice. Subjects were found not to be adept at dealing with data more than sufficient for problem solution; at screening out relevant from irrelevant data; or at employing indirect procedures. These are all skills that are demanded not only for the solving of code items but for the solving of problems in general. They are all skills that can in part be taught.

The research has dealt with only one type of problem situation, studying it in some detail. It is possible that the results are specific to the particular subjects and items and instructions used. The extent to which the findings can be generalised, both to other item types and to problem situations other than test items, can only be settled empirically, via further research.

****

REFERENCES

ABERCROMBIE, M. L. J. Psychological aids to education. Bull. Brit. Psychol. Soc., 1962, 49, 3-10.

ABORN, M. & RUBENSTEIN, H. Information theory and immediate recall. J. Exp. Psychol., 1952, 44, 260-66.

ADKINS, D. C. Construction and Analysis of Achievement Tests. Washington: Govt. Printing Office, 1947.

ANDERSON, N. S. & LEONARD, J. A. The recognition, naming, and reconstruction of visual figures as a function of contour redundancy. J. Exp. Psychol., 1958, 56, 262-70.

ARCHER, E. J., BOURNE, L. E. & BROWN, F. G. Concept identification as a function of irrelevant information and instructions. J. Exp. Psychol., 1955, 49, 153-64.

ATTNEAVE, F. Some informational aspects of visual perception. Psychol. Rev., 1954, 61, 183-93.

__________ Symmetry, information, and memory for patterns. Amer. J. Psychol., 1955, 68, 209-22.

BARTLETT, F. C. The Mind at Work and Play. London: Unwin, 1951.

__________ Thinking: an experimental and social study. London: Unwin, 1958.

BEAN, K. L. Construction of Educational and Personnel Tests. New York: McGraw-Hill, 1953.

BERLYNE, D. E. The influence of complexity and novelty in visual figures on orienting responses. J. Exp. Psychol., 1958, 55, 289-96.

BOURNE, L. E. & HAYGOOD, R. C. Effects of intermittent reinforcement of an irrelevant dimension and task complexity upon concept identification. J. Exp. Psychol., 1960, 60, 371-5.

__________ Supplementary report: effect of redundant relevant information upon the identification of concepts; J. Exp. Psychol., 1961, 61, 259-60.

__________ The role of stimulus redundancy in concept identification. J. Exp. Psychol., 1959, 58, 232-8.

BOURNE, L. E. & PENDLETON, R. B. Concept identification as a function of completeness and probability of information feedback. J. Exp. Psychol., 1958, 56, 413-420.

BREBNER, J. & GORDON, I. The influence of signal probability and the number of non-signal categories on selective reaction times. Quart. J. Exp. Psychol., 1964, 16, 56-60.

BRICKER, P. D. The identification of redundant stimulus patterns. J. Exp. Psychol., 1955, 49, 73-81.

BROADBENT, D. E. & HERON, A. Effects of a subsidiary task on performance involving immediate memory by younger and older men. Brit. J. Psychol., 1962, 53, 189-98.

BROWN, J. Some tests of the decay theory of immediate memory. Quart. J. Exp. Psychol., 1958, 10, 12-21.

BROWN, R. & LENNEBERG, E. A study in language and cognition in in perception. J. Abnorm. Soc. Psychol., 1954, 49, 454-62.

BRUNER, J. The Process of Education. Cambridge: Harvard University Press, 1960.

__________, GOODNOW, J. J. & AUSTIN, G. A. A Study of Thinking. New York: Wiley, 1956.

__________, & OLVER, R. Development of equivalence transformations in children. Monogr. Soc. Res. Child Develpm., 1963, 28, 125-41.

BULBROOK, M. E. An experimental inquiry into the existence and nature of "insight". Amer. J. Psychol., 1932, 44, 409-53.

BURACK, B. The nature and efficacy of methods of attack on reasoning problems. Psychol. Monogr., 1950, 64, no. 313.

BURT, C. The development of reasoning in school children. J. Exp. Ped., 1919, 5, 68-77 and 121-7.

__________ Mental and Scholastic Tests. London: King, 1921.

__________ The structure of the mind. A review of the results of factor analysis. Part II. Brit. J. Educ. Psychol., 1949, 19, 176-99.

__________ The theory of mental testing. Yearb. Educ., 1950, 42-73.

BUSWELL, G. T. Patterns of thinking in solving problems. University of California Publications in Education, 1956, 12, no. 2.

CALDER, J. R. Research in programmed learning in Aberdeen. Bull. Brit. Psychol. Soc., 1964, 17, p. 42.

CAMPBELL, A. C. Intelligence test items and serendipity. Aust. J. Psychol., 1963, 15, 201-5.

__________ Intelligence test items and serendipity: further considerations. Aust. J. Psychol., 1964, 16, in press.

CATTELL, R. B. A culture-free intelligence test I. J. Educ. Psychol., 1940, 31, 161-79.

__________ Personality and Motivation Structure and Measurement. London: Harrap, 1957.

CHANT, S. N. F. An objective experiment on reasoning. Amer. J. Psychol., 1933, 45, 282-91.

CHEN, H-P. Studies on Thurstone's Reasoning Tests. J. Genet. Psychol., 1937, 50, 15-25.

CHERRY, C. On Human Communication. New York: Science Editions, 1961.

CLAY, H. M. Changes of performance with age on similar tasks of varying complexity. Brit. J. Psychol., 1954, 45, 7-13.

COCHRAN, W. G. & COX, G. Experimental Designs. New York: Wiley, 2nd edn., 1957.

COHEN, J. Chance, Skill and Luck. Penguin Books, 1960.

CONRAD, R. Some effects on performance of changes in perceptual load. J. Exp. Psychol., 1955, 49, 313-22.

__________ Decay theory of immediate memory. Nature, 1957, 179, no. 4564.

__________ Accuracy of recall using Keyset and telephone dial, and the effect of a prefix digit. J. Appl. Psychol., 1958, 42, 285-8.

__________ Acoustic confusions in immediate memory. Quart. J. Exp. Psychol., 1964, 16, 75-84.

COOK, J. M., HEIM, A. W & WATTS, K. P. The Word-in-context: a new type of verbal reasoning test. Brit. J. Psychol., 1963, 54, 227-37.

CORMAN, B. R. The effect of varying amounts and kinds of information as guidance in problem solving. Psychol. Monogr., 1957, 71, no. 431.

DALE, A. B. The use of mental tests with university women students. Brit. J. Educ. Psychol., 1935, 5, 59-75.

DASHIELL, J. F. Direction orientation in maze running by the white rat. Compar. Psychol. Monogr., 1930, no. 32.

DODWELL, P. C. Some factors affecting the hearing of words presented dichotically. Canad. J. Psychol., 1964, 18, 72-91.

DONALDSON, M. C. The relevance to the theory of intelligence testing of the study of errors in thinking. Ph.D. thesis.

University of Edinburgh Library, 1956.

__________ Positive and negative information in matching problems. Brit. J. Psychol., 1959, 50, 253-62.

__________ A Study of Children's Thinking. In collaboration with Donald Withrington. London: Tavistock, 1963.

DOWNIE, N. M. Fundamentals of Measurement. London: Oxford University Press, 1958.

DRESSEL, P. L. The development of a test of uniqueness. Educ. Psychol. Measmt., 1957, 17, 43-54.

DUNCKER, K. On problem-solving. Psychol. Monogr. 1945, 58, no. 270. Translated by L. S. LEES.

DURKIN, H. E. Trial-and-error, gradual analysis, and sudden reorganisation: an experimental study of problem solving. Arch. Psychol. N. Y., 1937, no. 210.

EBEL, R. L. Writing the test item. In E. F. LINDQUIST (ed). Educational Measurement. Washington: American Council on Education, 1951, 185-249.

EDWARDS, A. L. Experimental Design in Psychological Research. New York: Holt, Rinehart & Winston, 1960.

ELITHORN, A., KERR, M. & MOTT, J. A group version of a perceptual maze test. Brit. J. Psychol., 1960, 51, 19-26.

ELITHORN, A., JONES, D., KERR, M. & LEE, D. The effects of the variation of two physical parameters on empirical difficulty in a perceptual maze test. Brit. J. Psychol., 1964, 55, 31-7.

ENTICKNAP, L. E. The perceptibility of the letters of an alphabet. Aust. J. Psychol., 1957, 9, 112-19.

ERIKSEN, C. W. Partitioning and saturation of visual displays and efficiency of visual search. J. Appl. Psychol., 1955, 39, 73-7.

FELDMAN, M. P. Motivation and task performance: a review of the literature. In H. J. EYSENCK (ed). Experiments in Motivation. London: Pergamon, 1964.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, A. & LEONARD, J. A. Stimulus correlates of visual pattern recognition. J. Exp. Psychol., 1956, 51, 1-11.

FITTS, P. D. & LEONARD, J. A. Stimulus Correlates of Visual Pattern Recognition. A probability approach. Columbus: Ohio State University, 1957.

FORGAYS, D. G. The development of differential word recognition. J. Exp. Psychol., 1953, 45, 165-8.

FURNEAUX, W. D. Intellectual abilities and problem-solving behaviour. In H. J. EYSENCK (ed). Handbook of Abnormal Psychology. Vol. II. London: Routledge & Kegan Paul, 1960, 167-92.

GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. & SPENCE, D. P. Cognitive control: a study of individual consistencies in cognitive behaviour. Psychol. Issues, 1959, 1, no. 4.

GARDNER, R. W. & LONG, R. I. Control, defence and centration effect: a study of scanning behaviour. Brit. J. Psychol., 1962(a), 53, 129-40.

____________ Cognitive controls of attention and inhibition: a study of individual consistencies. Brit. J. Psychol., 1962(b), 53, 381-8.

GARNER, W. R. Uncertainty and Structure as Psychological Concepts. New York: Wiley, 1962.

GEER, J. P. van de, A Psychological Study of Problem Solving.

Haarlem: Uitgeverij de Toorts, 1957.

GELFAND, S. Effects of prior associations and task complexity upon the identification of concepts. Psychol. Rep., 1958, 4, 567-74.

GOLDSTEIN, K & SCHEERER, M. Abstract and concrete behaviour. An experimental study with special tests. Psychol. Monogr., 1941, 53, no. 239.

GOTTSCHALDT, K. Der Aufbau des kindlichen Handelns. Zeitschr. Ang. Psychol. Beihefte, 1933, no. 68. Cited by R. S. WOODWORTH & H. SCHLOSBERG. Experimental Psychology. London: Methuen, 1954.

GREEN, B. F. & ANDERSON, L. K. Colour coding in a visual search task. J. Exp. Psychol., 1956, 51, 19-24.

GUILFORD, J. P., COMREY, A. L., GREEN, R. F. & CHRISTENSEN, P. R. A factor-analytic study of reasoning abilities. I. Hypotheses and description of tests. Reports from the Psychological Laboratory of the University of Southern California, 1950, no. 1.

GUILFORD, J. P. Figural and symbolic divergent-production abilities in adolescent and adult populations. Reports from the Psychological Laboratory of the University of Southern California, 1963, no. 29.

HANFMANN, E. A study of personal patterns in an intellectual performance. Character & Personality, 1941, 9, 315-25.

HANFMANN, E. & KASANIN, J. Conceptual thinking in schizophrenia. Nerv. Ment. Dis. Monogr., 1942, 67.

HARLOW, H. F. Learning sets and error factor theory. In KOCH, S. (ed). Psychology. A Study of a Science. Vol. II. New York: McGraw-Hill, 1959, 492-537.

HAYGOOD, R. C. & BOURNE, L. E. Forms of relevant stimulus redundancy in concept identification. J. Exp. Psychol., 1964, 67, 392-7.

HEBB, D. O. A Textbook of Psychology. London: Saunders, 1958.

HEIDBREDER, E. Towards a dynamic psychology of cognition. Psychol. Rev., 1945, 52, 1-22.

__________ The attainment of concepts: III. The process. J. Psychol., 1947, 24, 93-138.

__________ The attainment of concepts: VI. Exploratory experiments on conceptualisation at perceptual levels. J. Psychol., 1948, 26, 193-216.

HEIM, A. W. The Appraisal of Intelligence. London: Methuen, 1954.

HERMAN, D. T. & ENGSTRAND, R. Order effects in problem solving. Psychol. Rep., 1957, 3, 623-6.

HICK, W. E. On the rate of gain of information. Quart. J. Exp. Psychol., 1952, 4, 11-26.

HODGE, M. H. The influence of irrelevant information upon complex visual discrimination. J. Exp. Psychol., 1959, 57, 1-5.

HUNTER, I. M. L. The solving of five-letter anagram problems. Brit. J. Psychol., 1959, 50, 193-206.

__________ The organisation of problem solving. Inaugural Lecture. University of Keele. March, 1963.

HYMAN, R. Stimulus information as a determinant of reaction time. J. Exp. Psychol., 1953, 45, 188-96.

JAENSCH, E. R. Eidetic Imagery. London: Kegan Paul, 1930.

JEEVES, M. A. & DIENES, Z. P. Thinking in structures. Psychol. Monogr. on Cogn. Problems, 1964, 1, no. 1.

JOHN, E. R. Contributions to the study of the problem-solving process. Psychol. Monogr., 1957, 71, no. 447.

__________, & MILLER, J. G. The acquisition and application of information in the problem solving process. Behav. Sci., 1957, 2, 291-300.

JOHNSON, D. M. The Psychology of Thought and Judgment. New York: Harper, 1955.

__________, & HALL, E. R. Organisation of relevant and irrelevant words in the solution of verbal problems. J. Psychol., 1961, 52, 99-104.

KARP, S. A. Field dependence and overcoming embeddedness. J. Consult. Psychol., 1963, 27, 294-302.

KAY, H. Experimental studies of adult learning. Ph.D. thesis. University of Cambridge Library, 1953. Experiment described in A. T. WELFORD. Ageing and Human Skill. Oxford: University Press, 1958.

KESSEN, W. & KUHLMAN, C. (eds). Thought in the young child. Monogr. Soc. Res. Child Develpm., 1962, 27.

KOSSOV, B. B. Some methods for bringing out essential features in perceived objects. In Problems of Psychology. Oxford: Pergamon, 1960. Translated by R. G. HANKIN.

LINDLEY, E. H. A. A study of puzzles. Amer. J. Psychol., 1897, 8, 431-94.

LOEVINGER, J. Objective tests as instruments of psychological theory. Psychol. Rep., 1957, 3, 635-94.

LUCE, R. D. Response latencies and probabilities. In K. ARROW et al. (eds). Mathematical Methods in the Social Sciences 1959. Stanford: University Press, 1960, Ch. 20.

LUCHINS, A. S. Mechanisation in problem solving: the effect of Einstellung. Psychol. Monogr., 1942, 54, no. 248.

__________, & LUCHINS, E. H. New experimental attempts at preventing mechanisation in problem solving. J. Gen. Psychol., 1950, 42, 279-97.

MACCOBY, E. E. Developmental Psychology. Ann. Rev. Psychol., 1964, 15, 203-250.

McCOLSKEY, A. S. The qualitative investigation of intelligence. Brit. J. Psychol., 1957, 48, 276-85.

McGUIGAN, F. J. Incidental learning in the formation of concepts. Amer. J. Psychol., 1958, 71, 539-47.

McKELLAR, P. Imagination and Thinking. A psychological analysis. London: Cohen & West, 1957.

MACKINTOSH, N. J. The effect of irrelevant cues on reversal learning in the rat. Brit. J. Psychol., 1963, 54, 127-34.

MACKWORTH, N. H. Human problems of work design. Nature, 1949, 164, 982-3.

__________, & MACKWORTH, J. F. Visual search for successive decisions. Brit. J. Psychol., 1958, 49, 210-21.

MAIER, N. R. F. Reasoning in humans. I. On direction. J. Compar. Psychol., 1930, 10, 115-43.

__________ Reasoning in humans. II. The solution of a problem and its appearance in consciousness. J. Compar. Psychol., 1931, 12, 181-94.

__________ An aspect of human reasoning. Brit. J. Psychol., 1933, 24, 144-55.

MENGER, K. On the design of grouping problems and related intelligence tests. J. Educ. Psychol., 1953, 44, 275-87.

MILLER, G. A. Free recall of redundant strings of letters. J. Exp. Psychol., 1958, 56, 485-91.

MORGAN, J. J. Effects of non-rational factors on inductive reasoning. J. Exp. Psychol., 1944, 34, 159-68.

MOWBRAY, J. Choice reaction times for skilled responses. Quart. J. Exp. Psychol., 1960, 12, 193-202.

NEWELL, A., SHAW, J. C. & SIMON, H. A. Elements of a theory of human problem solving. Psychol. Rev., 1958, 65, 151-66.

______________ The processes of creative thinking. In H. E. GRUBER (ed). Contemporary Approaches to Creative Thinking. New York: Atherton Press, 1962, 63-119.

OSGOOD, C. E. A behaviouristic analysis of perception and language as cognitive phenomona. In J. S. BRUNER et al. (eds). Contemporary Approaches to Cognition. Cambridge: Harvard University Press, 1957, 75-118.

OSLER, S. F. & TRAUTMAN, G. E. Effect of stimulus complexity upon concept attainment at two levels of intelligence. J. Exp. Psychol., 1961, 62, 9-13.

PENROSE, L. S. & RAVEN, J. C. A new series of perceptual tests, preliminary communication. Brit. J. Med. Psychol.,1936, 16 97-104.

PETERSON, M. J. Some effects of the percentage of relevant cues and presentation methods on concept identification. J. Exp. Psychol., 1962, 64, 623-27.

PIAGET, J. Judgment and Reasoning in the Child. London: Kegan Paul, Trench & Trubner, 1928. Translated by M. WARDEN.

PILLINER, A. E. G. The measurement of educational achievement in Israel. U.N.E.S.C.O. Report, 1961, 1-97.

POLYA, G. How to Solve It. New York: Doubleday, 1957.

PORTEOUS, S. D. The Maze Test and Clinical Psychology. California: Pacific Books, 1959.

POSNER, M. I. An informational approach to thinking. University of Michigan Technical Report, April, 1962.

PRESSEY, S. L. Teaching machine (and learning theory) crisis. J. Appl. Psychol., 1963, 47, 1-9.

RAAHEIM, K. Problem solving and the ability to find replacements. Scand. J. Psychol., 1960, 1, 14-18.

REED, H. B. The learning and retention of concepts. IV. The influence of the complexity of the stimuli. J. Exp. Psychol., 1946, 36, 252-61.

RESTLE, F. A theory for discrimination learning. Psychol. Rev., 1955, 62, 11-19.

RIMLAND, B. Effect of extraneous numerical information on the U.S. Navy arithmetic test. Bureau of Naval Personnel, Technical Bulletin 59-4, March 1959.

RUBENSTEIN, H. & ABORN, M. Immediate recall as a function of degree of organisation and length of study period. J. Exp.Psychol., 1954, 48, 146-52.

SAUGSTAD, P. Problem-solving as dependent on availability of functions. Brit. J. Psychol., 1955, 46, 191-8.

SEMEONOFF, B. & TRIST, E. Diagnostic Performance Tests. London: Tavistock, 1958.

SIEGEL, S. Nonparametric Statistics. New York: McGraw-Hill, 1956.

SKINNER, B. F. Reply to Thouless. Aust. J. Psychol., 1963, 15, 92-3.

SMEDSLUND, J. Educational Psychology. Ann. Rev. Psychol., 1964, 15, 251-76.

SPEAKMAN, D. The effect of age on the incidental relearning of stamp values. J. Gerontol., 1954, 9, 162-7.

STANILAND, A. Redundancy as an experimental variable. Quart. J. Exp. Psychol., 1960, 12, 149-61.

STEVENSON, H. W. Piaget, behaviour theory, and intelligence. Monogr. Soc. Res. Child Developm., 1962, 27, 113-26.

STONE, G. C. & CALLAWAY, E. Effects of stimulus probability on reaction time in a number-naming task. Quart. J. Exp. Psychol., 1964, 16, 47-55.

SUTHERLAND, N. S. Stimulus analysing mechanisms. In Proceedings of a Symposium on the Mechanisation of Thought Processes. London: H.M.S.O., 1959, Vol. II, 575-609.

SZÉKELY, L. Knowledge and thinking. Acta Psychologica, 1950, 7, 1-24.

TAYLOR, C. W. (ed). Creativity: Progress and Potential. London: McGraw-Hill, 1964.

TAYLOR, D. W., HUNT, H. F. & GARNER, W. R. Education for research in psychology. Amer. Psychologist, 1958, 14, 167-79.

THORNDIKE, E. L. Human Learning. New York: Century, 1931.

THOULESS, R. H. The psychology teaching programme of Holland and Skinner. Aust. J. Psychol., 1963, 15, 85-91.

THURSTONE, L. L. The Nature of Intelligence. London: Kegan Paul, 1924.

__________ A Factorial Study of Perception. Chicago: University of Chicago Press, 1944.

UNDERWOOD, B. J. An orientation for research on thinking. Psychol. Rev., 1952, 59, 209-20.

WALLACE, J. G. Results of a test of high-grade intelligence applied to a university population. Brit. J. Psychol., 1952, 43, 61-9.

WALLAS, G. The Art of Thought. New York: Harcourt Brace, 1926.

WEAVER, H. E. & MADDEN, E. H. "Direction" in problem solving. J. Psychol., 1949, 27, 331-45.

WECHSLER, D. Cognitive, conative and non-intellective intelligence. Amer. Psychologist, 1950, 5, 78-83.

WELFORD, A. T. Ageing and Human Skill. London: Oxford University Press, 1958.

WERTHEIMER, M. Productive Thinking. New York: Harper, 1945.

WHITFIELD, J. W. An experiment in problem solving. Quart. J. Exp. Psychol., 1951, 3, 184-97.

WITKIN, H. A. Individual differences in ease of perception of embedded figures. J. Personality, 1950, 19, 1-15.

______________, DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R. & KARP, S. A. Psychological Differentiation. New York: Wiley, 1962.

______________, LEWIS, F. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B. & WAPNER, S. Personality through Perception. An experimental and clinical study. New York: Harper, 1954.

WOHLWILL, J. F. The teaching machine: psychology's new hobbyhorse. Teachers College Record, 1962, 64, 139-50.

______________ From perception to inference: a dimension of cognitive development. Monogr. Soc. Res. Child Develpm., 1962, 27, 87-107.

WOODWORTH, R. S. Experimental Psychology. London: Methuen, 1938.

______________ & SCHLOSBERG, H. Experimental Psychology. London: Methuen, 1954.

APPENDICES

APPENDIX I

TABLE OF CONTENTS

SAMPLE ITEMS FOR THE TESTS USED IN STUDY 1

SAMPLE ITEM WITH JUST SUFFICIENT DATA FOR SOLUTION

Example:

(1)CAPE (2)DISC (3)FUEL (4)SLIM

- X - -	.... (2)
- - - T	.... (1)
- - X -	.... (4)
- - T -	.... (3)

Each of the words (1)-(4) has been translated into a code, and the coded words are listed above. They are in a different order to their uncoded counterparts, and they are also incomplete.

The code is an arbitrary one, and to work out which incomplete coded word corresponds to which complete uncoded word it is necessary to consider the frequency of occurrence of particular letters and also the positions in which these letters occur.

Considering the above example, the second letter of the first coded word and the third letter of the third coded word are the same - both are X's. Hence these two words must be the coded versions of DISC and SLIM respectively, X standing for I.

The numbers (2) and (4) have therefore been written into the appropriate pairs of brackets.

Turning now to the remaining words, the second coded word ends with a T. T is also the third letter of the last coded word. Hence these two must be the coded versions of CAPE and FUEL respectively, with T standing for E.

The numbers (1) and (3) have therefore been written into the appropriate pairs of brackets.

Overleaf are code items 1-10 . Work through them in order.

You will be given 15 minutes to do as many of them as you can.

DO NOT TURN OVER UNTIL INSTRUCTED

SAMPLE ITEM WITH EXTRA IRRELEVANT LETTERS

Example:

(1)CAPE (2)DISC (3)FUEL (4)SLIM

R	X	-	-	............(2)
F	-	-	T	............(1)
B	-	X	-	............(4)
S	-	T	-	............(3)

Each of the words (1)-(4) has been translated into a code, and the coded words are listed above. They are in a different order to their uncoded counterparts, and they are also incomplete.

The code is an arbitrary one, and to work out which incomplete coded word corresponds to which complete uncoded word it is necessary to consider the frequency of occurrence of particular letters and also the positions in which these letters occur.

Considering the above example, the second letter of the first coded word and the third letter of the third coded word are the same - both are X's. Hence these two words must be the coded counterparts of DISC and SLIM respectively, X standing for I.

The numbers (2) and (4) have therefore been written into the appropriate pairs of brackets.

Turning now to the remaining words, the second coded word ends with a T. T is also the third letter of the last coded word. Hence these two must be the coded versions of CAPE and FUEL respectively, with T standing for E.

The numbers (1) and (3) have therefore been written into the appropriate pairs of brackets.

Overleaf are code items 1-10 . Work through them in order.

You will be given 15 minutes to do as many of them as you can.

DO NOT TURN OVER UNTIL INSTRUCTED

SAMPLE ITEM WITH FULL CODED DATA

Example:

(1)CAPE (2)DISC (3)FUEL (4)SLIM

R X B F ....(2)
F I L T ....(1)
B J X O ....(4)
S Z T J ....(3)

Each of the words (1)-(4) is listed above in code, the coded words being in a different order to their uncoded counterparts. The code is an arbitrary one, and to work out which coded word is which it is necessary to consider the frequency of occurrence of particular letters and also the positions in which these letters occur.

For example, the second letter of the first coded word and the third letter of the third coded word are identical - both are X's. Hence these two words must be the coded versions of DISC and SLIM respectively, X standing for I.

The numbers (2) and (4) have therefore been inserted in the appropriate pairs of brackets.

Turning now to the remaining words, the second coded word ends with T. T is also the third letter of the last coded word. Hence these two must be the coded versions of CAPE and FUEL respectively, with T standing for E.

The numbers (1) and (3) have therefore been inserted in the appropriate pairs of brackets.

Overleaf are code items 1-10 . Work through them in order.

You will be given 15 minutes to do as many of them as you can.

DO NOT TURN OVER UNTIL INSTRUCTED

SAMPLE ITEMS FOR THE TESTS USED IN STUDY 2

Example (i):

(1)LEARN (2)DAILY (3)ENVOY (4)REACH (5)DREAM

Q	W	J	P	G	....	( )
Q	P	T	Z	F	....	( )
J	B	K	S	F	....	( )
Z	J	P	W	B	....	( )
W	J	P	X	C	....	( )

Each of the words (1)-(5) has beer translated into a coded version, and these coded versions are listed above, in a different order to that of their uncoded counterparts. The code is completely arbitrary - i.e. it is not based on any rule such as replacing every letter by the letter third after it in alphabetical order. You are to discover which is the coded version of each of the words (1)-(5). This can be done by considering:

(a) the frequency of occurrence of different letters, and

(b) the positions in which they occur.

Considering Example (i), the first and second coded words both begin with a Q. Of the uncoded words, the only two with the same initial letter are DAILY and DREAM. Q must therefore stand for D. But is the first coded word DAILY and the second DREAM, or vice versa ?
Further scrutiny shows that these two coded words both contain a P - in fourth and second places respectively. From this it can be inferred that P stands for A, and that the first coded word is DREAM and the second DAILY. The numbers corresponding to these two words should now be placed in the appropriate pairs of brackets.

You may now notice that the second and third coded words both end in F. Since the second - Q P T Z F - is known to pepresent DAILY, F must stand for Y. The uncoded counterpart of the third coded word has therefore to be a word ending in Y. Only ENVOY fulfils this requirement, and its number should now be placed in the appropriate bmackets.
The two remaining words are REACH and LEARN. Since the third coded word - J B K S F - has been established as ENVOY, the letter B must stand for N. The coded version of LEARN must therefore end in a B. Hence the fourth coded word represents LEARN, and this leaves the fifth to be REACH. The numbers corresponding to these two words may now be placed in the appropriate pairs of brackets, and the item has been satisfactorily completed.

Example (ii):

(1)CAPE (2)DISC (3)FUEL (4)SLIM (5)CLAM

```
- X - -   ..............(   )
- - - T   ..............(   )
- - X -   ..............(   )
- - T -   ..............(   )
- - - -   ..............(   )
```

Here the coded versions of the words (1)-(5) are in an incomplete (skeleton) form. Sufficient letters are however given to enable an unambiguous solution of the problem, and it is tackled in the same sort of way as for Example (i).

Two of the skeleton coded words contain an X, and second and third places respectively. Looking at words (1)- 5), two possibilities arise:

(a) X stands for A, the first coded word being CAPE and the third CLAM
(b) X stands for I, the first coded word being DISC and the third SLIM

Another two of the skeleton coded words contain a T, in third and fourth places respectively. Only one possibility here presents itself: T stands for E, with the second coded word representing CAPE and the fourth FUEL. The numbers 1 and 3 may now be placed in the appropriate pairs of brackets.

The establishment of the second coded word as CAPE eliminates alternative (a) above. Hence X must stand for I, with the first coded word DISC and the third SLIM. By default, the fifth coded word must be CLAM - it is the only word left. The numbers for these three words should now be placed in the appropriate pairs of brackets, and the item has been satisfactorily completed.

ADDITIONAL SAMPLE ITEM DEMANDING THE USE OF INDIRECT PROCEDURES

- USED IN THE COURSE OF STUDY 3 -

Example (iii):

(1)WRAITH (2)MARKET (3)HAWSER (4)FAVOUR (5)CINDER

-	-	-	-	-	-	..............( )
-	-	G	-	B	P	..............( )
-	-	-	-	-	B	..............( )
-	J	-	-	-	-	..............( )
Y	-	-	-	-	-	..............( )

Here the coded version of the words (1)-(5) are again in an incomplete (skeleton) form, but this time there is only one pair of coded letters - the B's. The plan of attack for items of this sort is always to begin with the letters that are a pair. The B's occur in second last and last places and can only stand for T, the second coded word being WRAITH and the third MARKET. The numbers 1 and 2 may now be placed in the appropriate pairs of brackets. Since - - G - B P represents WRAITH, G stands for A and P stands for H. As for the other two coded letters - the J and the Y - what they stand for is not known. It is however known that they do not stand for A or H (or T). Hence the fourth coded word represents a word that does not have an A or a H or a T in second position. Of the as yet unidentified words - HAWSER, FAVOUR and CINDER - only CINDER fulfils this requirement. Hence the fourth coded word represents CINDER, J stands for I, and the number 5 may now be placed in the appropriate brackets.

Take now the fifth coded word. It begins with a Y. Whatever Y stands for (which is not known), it does not stand for A or H or T or I. Hence the fifth coded word cannor represent a word beginning with A or H or T or I. This eliminates HAWSER. Hence the fifth coded word is FAVOUR and the first coded word must then represent HAWSER, since it is the only word left. The numbers 4 and 3 may now be inserted in the appropriate pairs of brackets, and the item has been satisfactorily completed.

TESTS USED IN STUDY 4

(Solution of the examples gone through in like manner to that for the sample items used in Studies 1 and 2)

CODES TEST I

NAME: ____

Example I:

(1)PRISM (2)FLECK (3)TENOR (4)FROST (5)CRIME

Y	-	-	-	-	..............( )
Z	-	-	-	Y	..............( )
Z	-	-	J	-	..............( )
-	-	-	-	-	..............( )
J	-	-	-	-	..............( )

Example II:

(1)EVENT (2)NEVER (3)TREND (4)DROWN (5)BOUND

-	S	-	-	-	..............( )
S	-	-	-	H	..............( )
H	-	-	-	M	..............( )
-	-	-	-	M	..............( )
-	-	-	-	-	..............( )

DO NOT TURN OVER UNTIL TOLD

Just sufficient data and no false leads

I. (1)HELPS (2)CHOPS (3)CIDER (4)WREAK (5)FROWN

- X - - - ....(1)
- B X - - ....(4)
- - - - - ....(3)
- B G - - ....(5)
- - G - - ....(2)

II. (1)OVERT (2)STAIR (3)SHAME (4)APRIL (5)FRESH

- - Z - - ....(4)
- - Q Z - ....(1)
J - - - Q ....(3)
- - - - - ....(5)
J - - - - ....(2)

III. (1)ROUSE (2)ANKLE (3)KILNS (4)SNORE (5)SLICK

D T - - - ....(3)
- - D - - ....(2)
- - - - - ....(1)
X - - - - ....(4)
X - T - - ....(5)

IV. (1)EVICT (2)TEMPT (3)LARVA (4)MOTOR (5)CREAM

H - - - G ....(5)
- - - - - ....(2)
- F - H - ....(1)
G - - - - ....(4)
- - - F - ....(3)

GO STRAIGHT ON

V. (1)TORCH (2)FRESH (3)ROAST (4)SNORE (5)CLAMP

J - P - - ....(3)
- J - - - ....(2)
- - - Z - ....(1)
- - - - - ....(4)
Z - P - - ....(5)

VI. (1)TOWER (2)WATER (3)SPAWN (4)SPRAT (5)GRIND

X - - - - ....(1)
- - Q B X ....(4)
- Q - - - ....(5)
- - B - - ....(3)
- - - - - ....(2)

VII. (1)GRAIN (2)ASIDE (3)PYLON (4)GULPS (5)EAGLE

- - - - - ....(3)
- J - - - ....(2)
V - T - J ....(4)
V - - - - ....(1)
- - - T - ....(5)

VIII. (1)EARTH (2)MONTH (3)PRISM (4)EMITS (5)PLUMS

D C - - - ....(3)
- - C - - ....(1)
- - - - W ....(2)
- - - - - ....(4)
D - - - W ....(5)

GO STRAIGHT ON

IX. (1)FROST (2)DRIFT (3)AFIRE (4)TARTS (5)RADIO

- - L - U ....(1)
- K - - - ....(2)
- - - - L ....(5)
- - - - - ....(3)
- - K U - ....(4)

X. (1)MONEY (2)SHUTS (3)PENNY (4)SPORT (5)MINUS

- Z - - - ....(3)
X - - Z - ....(1)
X - - W - ....(5)
- - W - - ....(2)
- - - - - ....(4)

XI. (1)RACES (2)WRECK (3)IMBUE (4)SCREW (5)WASTE

- - - P M ....(4)4
Y - - - - ....(1)
M - - - - ....(5)
- - - - - ....(3)
- Y P - - ....(2)

XII. (1)CREAM (2)PEACH (3)DREAM (4)HEARD (5)HAREM

O K - - - ....(1)
- - - O - ....(2)
S - K - - ....(5)
S - - - - ....(4)
- - - - - ....(3)

CODES TEST II

NAME: ____

Example I:

(1)PRISM (2)FLECK (3)TENOR (4)FROST (5)CRIME

Y	-	-	-	-	....	( )
Z	-	-	-	Y	....	( )
Z	-	-	J	-	....	( )
-	-	-	-	-	....	( )
J	-	-	-	-	....	( )

Example II:

(1)EVENT (2)NEVER (3)TREND (4)DROWN (5)ROUND

-	-	-	-	-	....	( )
-	-	-	F	-	....	( )
-	S	-	-	-	....	( )
-	X	-	F	-	....	( )
-	S	X	-	-	....	( )

DO NOT TURN OVER UNTIL TOLD

Just sufficient data with false leads present

I. (1)HARPS (2)WRATH (3)FROWN (4)SHOPS (5)SPEAR

U X – Q I ....( )
– B X – U ....( )
I – – – – ....( )
– B G – – ....( )
– – G Q – ....( )

II. (1)OVERT (2)FRESH (3)APRIL (4)STAIR (5)SLIME

– – Z – X ....( )
– – Q Z C ....( )
J X D – Q ....( )
– – – – – ....( )
J C – D – ....( )

III. (1)ANKLE (2)SLICK (3)SNARE (4)KILNS (5)RAISE

D T – P – ....( )
– P D – – ....( )
– W – – H ....( )
X – W – H ....( )
X – T – – ....( )

IV. (1)CREAM (2)TEMPT (3)CARVE (4)METER (5)EVICT

H J K – G ....( )
L K – – – ....( )
– F – H L ....( )
G – – – – ....( )
– – J F – ....( )

GO STRAIGHT ON

V. (1)TORCH (2)FRESH (3)ROAST (4)SCORE (5)CRASH

J	-	P	-	-	....	( )
-	J	-	-	-	....	( )
-	-	-	Z	-	....	( )
-	-	-	-	-	....	( )
Z	-	P	-	-	....	( )

VI. (1)TOWER (2)WANTS (3)PRAWN (4)SPRAT (5)TREND

X	-	-	-	-	....	( )
-	-	Q	B	X	....	( )
-	Q	-	-	-	....	( )
-	-	B	-	-	....	( )
-	-	-	-	-	....	( )

VII. (1)GRAIN (2)ASIDE (3)PYLON (4)GULPS (5)ANGLE

-	-	-	-	-	....	( )
-	J	-	-	-	....	( )
V	-	T	-	J	....	( )
V	-	-	-	-	....	( )
-	-	-	T	-	....	( )

VIII. (1)EARTH (2)MONTH (3)PRISM (4)EMITS (5)PLANS

D	C	-	-	-	....	( )
-	-	C	-	-	....	( )
-	-	-	-	W	....	( )
-	-	-	-	-	....	( )
D	-	-	-	W	....	( )

GO STRAIGHT ON

IX. (1)FROST (2)TARTS (3)RADIO (4)DRAFT (5)AFTER

J - L N U ....( )

- K M - - ....( )

- M - - L ....( )

- J - - - ....( )

- - K U N ....( )

X. (1)MONEY (2)SHEET (3)MINOR (4)PENNY (5)SPORT

- Z - - B ....( )

X - - Z B ....( )

X - - W - ....( )

A - W - C ....( )

A - - - C ....( )

XI. (1)WATER (2)RACES (3)SCREW (4)ACTOR (5)WRECK

J - - P M ....( )

Y - B - J ....( )

M U - - - ....( )

U - - - - ....( )

- Y P B - ....( )

XII. (1)HAREM (2)REACH (3)DREAM (4)HEARD (5)CREAM

O K - L - ....( )

- - L O - ....( )

S - K - B ....( )

S W - - - ....( )

- - W - B ....( )

CODES TEST III

NAME: ____

Example I:

(1)PRISM (2)FLECK (3)TENOR (4)FROST (5)CRIME

Y	A	-	-	B	....	( )
Z	B	-	-	Y	....	( )
Z	-	-	J	-	....	( )
-	-	F	-	-	....	( )
J	-	F	-	A	....	( )

Example II:

(1)EVENT (2)NEVER (3)TREND (4)DROWN (5)ROUND

-	S	-	-	A	....	( )
S	-	-	-	H	....	( )
H	A	-	-	-	....	( )
-	X	-	F	M	....	( )
M	-	X	-	F	....	( )

DO NOT TURN OVER UNTIL TOLD

Redundant data with false leads present

I. (1)HARPS (2)WRATH (3)FROWN (4)SHOPS (5)SPEAR

U	X	–	Q	I	...............(	)
–	B	X	–	U	...............(	)
I	–	–	–	–	...............(	)
–	B	G	–	–	...............(	)
–	–	G	Q	–	...............(	)

II. (1)OVERT (2)FRESH (3)APRIL (4)STAIR (5)SLIME

–	–	Z	–	X	...............(	)
–	–	Q	Z	C	...............(	)
J	X	D	–	Q	...............(	)
–	–	–	–	–	...............(	)
J	C	–	D	–	...............(	)

III. (1)ANKLE (2)SLICK (3)SNARE (4)KILNS (5)RAISE

D	T	–	P	–	...............(	)
–	P	D	–	–	...............(	)
–	W	–	–	H	...............(	)
X	–	W	–	H	...............(	)
X	–	T	–	–	...............(	)

IV. (1)CREAM (2)TEMPT (3)CARVE (4)METER (5)EVICT

H	J	K	–	G	...............(	)
L	K	–	–	–	...............(	)
–	F	–	H	L	...............(	)
G	–	–	–	–	...............(	)
–	–	J	F	–	...............(	)

GO STRAIGHT ON

V. (1)FRESH (2)SCORE (3)TORCH (4)CRASH (5)ROAST

J	–	P	B	–	................( )
–	J	–	–	N	................( )
–	Q	–	Z	–	................( )
B	–	Q	–	–	................( )
Z	–	P	–	N	................( )

VI. (1)TOWER (2)WANTS (3)TREND (4)SPRAT (5)PRAWN

X	–	M	J	–	................( )
C	–	Q	B	X	................( )
–	Q	J	–	–	................( )
–	–	B	M	–	................( )
–	–	–	–	C	................( )

VII. (1)ANGLE (2)GRAIN (3)ASIDE (4)PYLON (5)GULPS

–	–	–	–	H	................( )
–	J	Z	–	–	................( )
V	–	T	–	J	................( )
V	–	M	Z	H	................( )
M	–	–	T	–	................( )

VIII. (1)PRISM (2)EMITS (3)MONTH (4)EARTH (5)PLANS

D	C	–	–	–	................( )
–	E	C	–	U	................( )
–	–	–	J	W	................( )
–	–	–	J	U	................( )
D	–	B	–	W	................( )

GO STRAIGHT ON

IX. (1)FROST (2)TARTS (3)RADIO (4)DRAFT (5)AFTER

J	–	L	N	U	................( )
–	K	M	–	–	................( )
–	M	–	–	L	................( )
–	J	–	–	–	................( )
–	–	K	U	N	................( )

X. (1)MONEY (2)SHEET (3)MINOR (4)PENNY (5)SPORT

–	Z	–	–	B	................( )
X	–	–	Z	B	................( )
X	–	–	W	–	................( )
A	–	W	–	C	................( )
A	–	–	–	C	................( )

XI. (1)WATER (2)RACES (3)SCREW (4)ACTOR (5)WRECK

J	–	–	P	M	................( )
Y	–	B	–	J	................( )
M	U	–	–	–	................( )
U	–	–	–	–	................( )
–	Y	P	B	–	................( )

XII. (1)HAREM (2)REACH (3)DREAM (4)HEARD (5)CREAM

O	K	–	L	–	................( )
–	–	L	O	–	................( )
S	–	K	–	B	................( )
S	W	–	–	–	................( )
–	–	W	–	B	................( )

Set I items

(1)MOIST (2)BLAST (3)DRIVE (4)FREAK (5)CHEER

–	J	W	–	–	................( )
–	–	–	Z	–	................( )
–	–	W	Z	–	................( )
–	J	–	–	–	................( )
–	–	–	–	–	................( )

(1)POKER (2)APART (3)SLUMS (4)BAKER (5)GRANT

–	–	–	–	I	................( )
–	D	–	–	–	................( )
–	I	D	–	Q	................( )
–	–	–	–	–	................( )
–	–	–	–	Q	................( )

(1)TRIES (2)VESTS (3)BELIE (4)JOLLY (5)CREAK

–	–	F	–	X	................( )
–	–	–	–	–	................( )
–	–	F	–	–	................( )
–	G	–	–	–	................( )
–	G	X	–	–	................( )

(1)ARGUE (2)CHIME (3)GLAND (4)SQUID (5)GRUEL

–	–	F	–	–	................( )
F	–	Z	–	P	................( )
–	–	–	–	–	................( )
–	–	–	–	P	................( )
Z	–	–	–	–	................( )

Set II items

(1)FEAST (2)GREAT (3)MOIST (4)INERT (5)PRINT

- J W - B ....( )
- - - Z B ....( )
- - W Z - ....( )
- J - - B ....( )
- - - - B ....( )

(1)COVER (2)CLAMP (3)CRAMP (4)CATER (5)CHURN

F - - - I ....( )
F D - - - ....( )
F I D - Q ....( )
- - - - - ....( )
F - - - Q ....( )

(1)DREAM (2)DRIER (3)DERBY (4)DELVE (5)DOLLS

N - F - X ....( )
N - - - - ....( )
N - F - - ....( )
N G - - - ....( )
- G X - - ....( )

(1)ALBUM (2)GLOOM (3)SLICE (4)BLADE (5)ALIGN

- T F - - ....( )
F - Z - P ....( )
- T - - - ....( )
- T - - P ....( )
Z T - - - ....( )

Set III items

(1)CLASP (2)DREAM (3)HOIST (4)QUELL (5)BRIDE

–	J	W	–	–	....	( )
G	–	–	Z	F	....	( )
–	–	W	Z	–	....	( )
–	J	–	–	Y	....	( )
N	–	–	–	–	....	( )

(1)ASHEN (2)MOTOR (3)SATYR (4)PLANT (5)CRAFT

–	–	–	B	I	....	( )
U	D	–	–	–	....	( )
–	I	D	–	Q	....	( )
[illegible]	–	–	G	–	....	( )
K	–	–	–	Q	....	( )

(1)BREAK (2)PRINT (3)KETCH (4)DELVE (5)SOLID

–	–	F	–	X	....	( )
–	–	–	M	–	....	( )
J	–	F	–	–	....	( )
–	G	–	–	–	....	( )
Y	G	X	U	–	....	( )

(1)GRAIN (2)PYLON (3)CHUTE (4)ANGLE (5)GULPS

–	B	F	Q	–	....	( )
F	–	Z	–	P	....	( )
–	M	–	W	–	....	( )
–	–	–	–	P	....	( )
Z	–	–	–	–	....	( )

CODES TEST

NAME: ______________________________

EXAMPLE I:

(1)PRISM (2)FLECK (3)TENOR (4)FROST (5)CRIME

Y	-	-	-	-	..............( )
Z	-	-	-	Y	..............( )
Z	-	-	J	-	..............( )
-	-	-	-	-	..............( )
J	-	-	-	-	..............( )

EXAMPLE II:

(1)DREAD (2)BRAVE (3)BRICK (4)CRATE (5)TRAIN

J	F	-	-	-	..............( )
-	F	-	J	M	..............( )
-	F	S	-	-	..............( )
-	-	M	S	-	..............( )
-	F	-	-	-	..............( )

DO NOT TURN OVER UNTIL TOLD

Sample items used in conjunction with the items from

Sets I and II

CODES TEST

NAME: ____

EXAMPLE I:

(1)TENOR (2)PRISM (3)FLECK (4)FROST (5)CRIME

Y	-	-	-	-	....	( )
Z	-	-	-	Y	....	( )
Z	-	-	J	-	....	( )
-	-	-	-	-	....	( )
J	-	-	-	-	....	( )

EXAMPLE II:

(1)MAULS (2)FLOWN (3)QUITE (4)BOARD (5)TRACK

-	J	Y	-	-	....	( )
V	-	Y	J	-	....	( )
-	H	-	-	X	....	( )
G	-	-	H	-	....	( )
-	-	-	-	B	....	( )

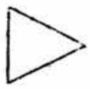
DO NOT TURN OVER UNTIL TOLD

Sample items used in conjunction with the items of

Sets II and III

Embedded Figures Test

NAME: ____________________________

Look at this sample figure: 

The sample figure is contained in each of the figures below. It is in the same orientation and of the same size as in the sample. Figure (1) has baen marked to show this.

Find where the sample figure is in Figure (2) and mark it.

Figure (1)

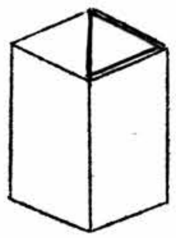
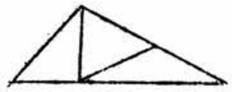


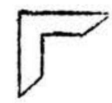
Figure (2)



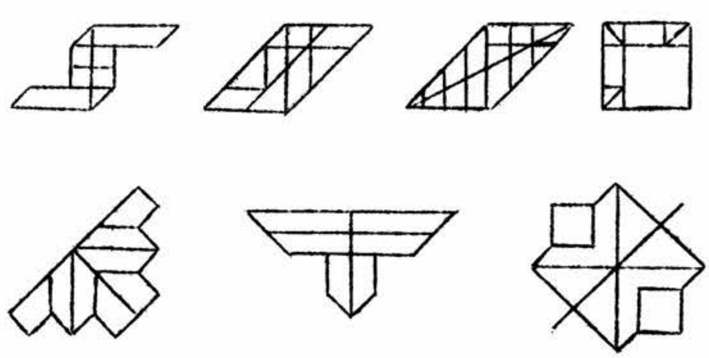
DO NOT TURN OVER UNTIL TOLD

PART I

Look at the adjacent figure.



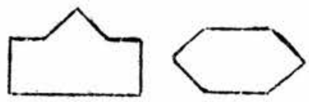
This figure is contained in each of the drawings below. Find where it is in each drawing and then mark it. Mark only one figure in each drawing.



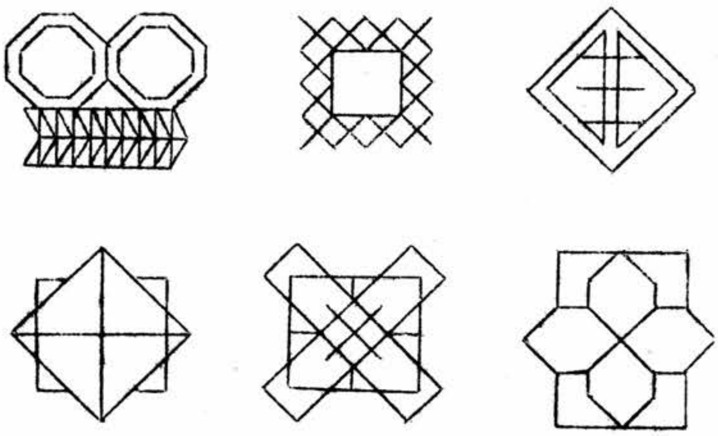
DO NOT TURN OVER UNTIL TOLD

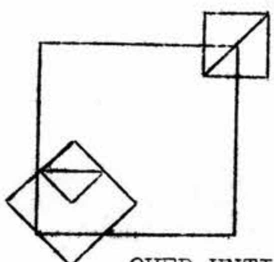
PART II.

Look at the two adjacent figures.



One of these figures is contained in each of the drawings below. In each of the drawings mark that part which is the same as one or other of the adjacent figures. Mark only one figure in each drawing.

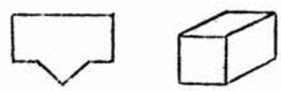




DO NOT TURN OVER UNTIL TOLD

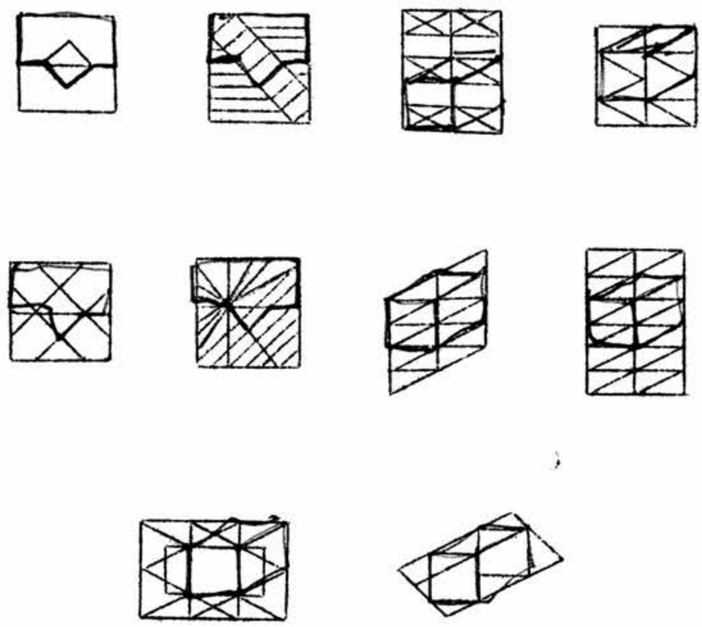
PART III

Look at the two adjacent figures.



One of them is contained in each of the drawings below.

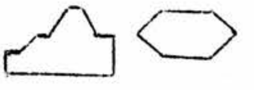
In each of the following drawings mark that part which is the same as one of the above figures. Mark only one figure in each drawing.



DO NOT TURN OVER UNTIL TOLD

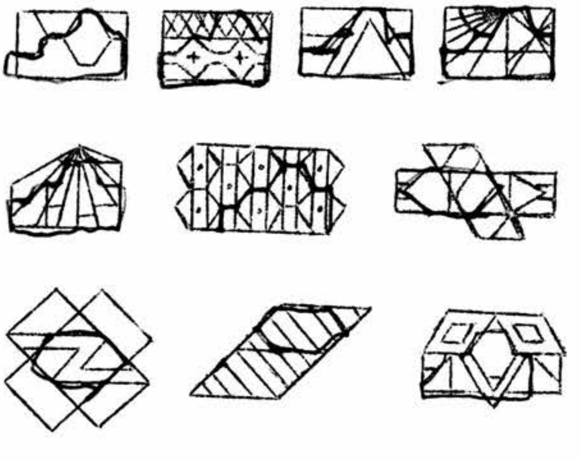
PART IV

Look at the two adjacent figures.



One of them is contained in each of the drawings below.

In each of the following drawings mark that part which is the same as one of the above figures. Mark only one figure in each drawing.



SAMPLE ITEMS USED IN STUDY 5

Example I: (redundant)

T O A S T	L - V - F	____________
M A N O R	- B F L -	____________
C R I M E	- H - J B	____________
P R I S M	- J H - -	____________
T R E C K	- - V - -	____________

Example II: (nonredundant)

C H A I N	- X - - -	____________
D R I V E	- - B X -	____________
H O V E L	M - - - -	____________
C H O R E	- B - - M	____________
L O A D S	- - - - -	____________

Example III: (redundant)

M O T O R	H - - B G	____________
E V I C T	- F - H D	____________
T E M P T	D - - - -	____________
C R E A M	G - - - J	____________
L A R V A	- B J F -	____________

Example IV: (nonredundant)

N E V E R	- X - - -	____________
T R E N D	- - X - F	____________
D R O W N	- S - - -	____________
E V E N T	S - - F -	____________
R O U N D	- - - - -	____________

INSTRUCTIONS USED IN CONJUNCTION WITH THE SAMPLE ITEMS

FOR STUDY 5

(These may be taken as illustrative of the explanations and instructions used for all the studies involving individual testing)

Example I:

Here are five words listed down the page - **TOAST** MANOR CRIME PRISM TRECK. Next to them are listed coded versions of the words. Not all the coded letters have been filled in. Those that have been filled in occur in pairs, e.g. a pair of L's, a pair of V's, and so on. Enough letters have been filled in for you to work out which incomplete coded word represents each of the words TOAST MANOR CRIME PRISM TRECK. To work this out you take the various letter pairs and note the positions in which they occur.

Here is an L - in first position. Here is another L - in fourth position. The next step is to look through this list of words for a pair of words, one of them having a certain letter in first position while the other has that same letter in the fourth position.

The first word - TOAST - begins with a T. Does any word have a T in fourth position ? .....No..... Therefore L does not stand for T.

The second word - MANOR - begins with an M. CRIME has an M in fourth position. Hence the first coded word could be MANOR and the second CRIME. But further inspection shows that these two words do not, in fact, fit in with the pattern of coded letters. If the first coded word were MANOR then F would stand for R, while if the second coded word were CRIME then F would stand for I. But F cannot stand for two different letters. Each coded letter stands for one letter and one letter only, in the uncoded versions. Hence the hypothesis that the first coded word is MANOR and the second CRIME must be rejected. Even if you had not noted that the F's fail to fit in with this hypothesis, you would soon have found out that the hypothesis is wrong. Thus if the first coded word (L - V - F) represented MANOR then V would stand for N. The fifth coded word would then represent a word with N in the middle position. But no such word is given.

Suppose now that L stands for C. This would make the first coded word CRIME and the second TRECK. If L - V - F represents CRIME then V stands for I. This being so, the fifth coded word could be PRISM. And if the second coded word (- B F L -) represents TRECK then B stands for R. The third coded word will then represent a word ending in R, and MANOR fits in with this. Only the fourth coded word now remains unaccounted for: to discover its identity you do not need to look at any coded letter. It must be TOAST, since that is the only word left. If you check on the various letter pairs, you will see that they do all in fact fit the pattern.

P.T.O.

All the items in the test proper are basically similar to the example we have just done.

Always there is only one ordering of the words that is perfectly correct - i.e. that is completely consistent with the pattern of coded letters.

Look now at Example II, which is on the next page of your booklet. Here there are not as many letter pairs as in Example I, but there are still enough to arrive at an unambiguous solution. Suppose we begin with the X's. They occur in second and fourth positions, so we need to find a pair of words one of which has a certain letter in second position while the other has this same letter in fourth position. DRIVE and CHORE fit the bill. But are they the sole possibilities? ..... Yes ..... Hence the first coded word is DRIVE and the second CHORE. Since - - B X - represents CHORE, B stands for O. There is another B in the second position in the fourth coded word. Hence the fourth coded word must represent a word with O in second position. Two possibilities arise: HOVEL and LOADS. Which of these is correct?

Suppose the fourth coded word is LOADS. Then M would stand for S. The third coded word would then represent a word beginning with S. But no such word is listed. Hence the fourth coded word cannot represent LOADS and must instead represent HOVEL. This being so, M stands for L. The third coded word will then be a word beginning with L - LOADS. The fifth coded words must then be CHAIN, since it is the only word left.

There are two more examples overleaf, which you can do for yourself. You can use the margins of the pages for scribbling paper as much as you like. I will tell you when to turn over and start each item, and, so that I can get the time score correct, will you tell me when you have finished each item to your satisfaction.

Any questions ?

When I say "go", turn over and start Example 3.

"Go".

..... etc.

The above gives some indication of the way in which the items were presented to the subjects who were individually tested. The actual wording was not rigidly standardised, it being considered that the important thing is to ensure that the subject understands what is required of him, rather than for the tester to utter a fixed set of verbal statements. If a subject sought elucidation of the procedure at any point, this was given. Some subjects took over the solution of the examples for themselves, right from the start, rather than sitting back and letting the tester demonstrate the solutions to them. When this happened care was taken to ensure that the subject was aware that plausible-but-incorrect alternative identifications can at times arise for a given letter pair.

SAMPLE ITEMS USED IN STUDY 6

Example I:

P L E B S	- - - X -	______________
F O C U S	- X G - T	______________
M A R K S	- Z - - -	______________
A R I E S	- - - - T	______________
H U M P S	G - Z - -	______________

Example II:

S P O T S	- E B N -	______________
S O C K S	J - - Z N	______________
S M A R T	- B - - -	______________
S U R L Y	- - Z - -	______________
S H I P S	J - - E -	______________

Example III:

P R I C E	V - - J -	______________
G R A N T	- - D - -	______________
C R U S T	D - - - J	______________
B R I N K	- M - - -	______________
A R E A S	- M - V -	______________

Example IV:

O X L I P	Z C - J N	______________
F O L L Y	- H Q - -	______________
P E L T S	K N C J	______________
I S L E T	H - - K Z	______________
B A L M Y	- - Q - -	______________

For use with the P = 0.00 items

Example I:

S K I R T	- B - - -	________
S A C K S	G D - - -	________
S U R L Y	- - - D -	________
S P O I L	G X - - -	________
S O U P S	- - X B -	________

Example II:

I R A T E	P - - N -	________
P R I S M	- - - F -	________
C R U M B	- - - - N	________
B R A K E	- - - G -	________
T R I C K	F - - P G	________

Example III:

M O U N D	B I - - -	________
D A R N S	- - - - B	________
A L O N E	- - - K -	________
L E A N S	- F - - -	________
S H U N T	F - I K -	________

Example IV:

S I N C E	D K L - -	________
A G A T E	- Q M - J	________
U N I T E	L D - - J	________
R O U T E	K - Y - -	________
G R A V E	Y M Q - -	________

For use with the P = 0.50 items

SAMPLE ITEMS USED IN STUDY 7

<u>Example I</u>:

C O A S T	F - - - -	______________
D R I V E	- - K - M	______________
B O U N D	F - K - -	______________
C R O P S	- - - - M	______________
S T A G E	- - - - -	______________

<u>Example II</u>:

C L A I M	B X - - -	______________
I C O N S	- - - T -	______________
R A I L S	X - T - -	______________
M I L E S	- - - - -	______________
A R E A S	- B - - -	______________

<u>Example III</u>:

M I L E S	H - Q Z -	______________
B L I N D	H - - - -	______________
M O I S T	- - - - -	______________
C L A S P	- - Q - -	______________
G R A N D	- - - Z -	______________

<u>Example IV</u>:

B E L O W	- - I - Q	______________
S O U N D	- I - - -	______________
C A R D S	- - - - -	______________
B L E S T	- - - C -	______________
D R O W N	- - - Q C	______________

SAMPLE ITEMS USED IN STUDY 8

Example I: (P = 0.00)

S E R V E	H - - B -	________________
T I N G E	- - - - D	________________
C L O V E	- - W - -	________________
A R G U E	- - - H D	________________
S L A T E	- W B - -	________________

Example II: (P = 0.50)

M O U N D	B I - - -	________________
D A R N S	- G - - B	________________
A L O N E	Z - - - -	________________
L E A N S	- F G - -	________________
S H U N T	F - I - Z	________________

Example III: (P = 0.50)

S K I R T	- B F - -	________________
S A C K S	G D - N -	________________
S U R L Y	- - - D -	________________
S P O I L	G X N - -	________________
S O U P S	- F X B -	________________

Example IV: (P = 0.75)

H A R E M	- Z - F -	________________
P E R M S	- - - - G	________________
E A R L Y	Z - - G -	________________
A P R I L	- - - - -	________________
C U R S E	- - - - F	________________

SAMPLE ITEMS USED IN STUDY 9

Example I: (just sufficient data)

C L I M B	F - - - H	________
U N T I E	X - - F -	________
B O N D S	G - - - -	________
B A T O N	- - - - -	________
T R A C K	- - Q - -	________

Example II: (just sufficient data)

F O C U S	- - Z D -	________
C U R L S	- - - - Q	________
L O A N S	- K - G -	________
S H U T S	- - - - -	________
B L E A T	- Z - V -	________

Example III: (extra irrelevant letters)

B R A V E	F - - - -	________
S W A M P	- G N - I	________
S T O V E	Q - X - G	________
C L A S P	- - - U -	________
H O V E L	- - T - -	________

APPENDIX II

APPENDIX II

TABLE OF CONTENTS

APPENDIX II

Report of a pilot study on the effects of irrelevant data on the solving of number series

Suppose one takes a simple series of numbers such as 2 4 6 8 and introduces an irrelevant number thereto, e.g. - 2 4 6 7 8. The subject can be asked to locate the number that violates the rule of construction for the series, and the item has become a number series correction item. This is a simple example, and it is obvious which number is the odd-man-out. Take however the item $1\frac{3}{4}$ $2\frac{1}{2}$ $2\frac{3}{4}$ $3\frac{1}{2}$ $4\frac{3}{4}$ $6\frac{1}{4}$ 8. Here it is not immediately apparent which is the odd-man-out. None of five undergraduates to whom this item was shown just to try it out solved it without considerable delay. The following eight number series correction items were constructed:

(1)	11	13	9	8	11	5	9	
(2)	4	8	10	12	20	22	44	
(3)	11	18	22	29	58	65		
(4)	64	32	16	12	3	-1	$-\frac{1}{4}$	
(5)	$1\frac{3}{4}$	$2\frac{1}{2}$	$2\frac{3}{4}$	$3\frac{1}{2}$	$4\frac{3}{4}$	$6\frac{1}{4}$	8	
(6)	48	144	40	90	120	32	96	
(7)	50	64	30	8	32	36	6	18
(8)	2	3	5	8	11	13	21	

In none of these items could the number irrelevant to solution, which has been underlined in each case, be described as "obviously irrelevant": it is in each case akin to the relevant numbers, and there is no ready basis by which it may be distinguished therefrom. Subjects have to locate the irrelevant number, and then fill in the two numbers at the beginning of the item as proof of having grasped the structural principle of the series. It is stated in the instructions that the irrelevant number is never the first or last of the series of numbers given. The eight items above were administered to 28 Second Ordinary psychology students at the University of Edinburgh. A comparable group of 25 students was given the same eight items without the irrelevant numbers. There were two sample items in each case, and 15 minutes working time was allowed. 2 points were allotted per item, making a maximum possible score of 16. It was predicted that items (1)-(8), in which irrelevant numbers are present, will be found more difficult than the same items without the irrelevant numbers, due to the irrelevant numbers not being "obviously irrelevant" and being therefore distracting and misleading. Mean scores were as follows:

Items without an irrelevant number (N=25)	Items with an irrelevant number (N=28)
$\bar{X}$ = 10.6	$\bar{X}$ = 3.4
(R = 4-16)	(R = 0-10)

The difference between groups is highly significant ($P < 0.0001$, 1-tailed; Mann-Whitney U test). As predicted, the presence of an irrelevant number seriously impedes solution. The subjects' spontaneous comments on the items and their reported feelings of frustration and helplessness further confirm that the difficulty caused by the irrelevant numbers is overwhelming. One reason for this is that the subjects were not given any indication as to the sorts of rules on which the series are based. Another reason is that because the irrelevant numbers are not "obviously irrelevant" they may suggest alternative solutions, initially plausible, but not in fact correct. There is the possibility of this happening with item (5) for example. Working from left to right - as most subjects do - the first four numbers ($1\frac{3}{4}$ $2\frac{1}{2}$ $2\frac{3}{4}$ $3\frac{1}{2}$) may suggest that the rule for the series is to add $\frac{3}{4}$ and $\frac{1}{4}$ alternately to the preceding number, but this rule breaks down at the next step. The number irrelevant to the series is in fact the third ($2\frac{3}{4}$), the others forming a series in which the difference between each pair of numbers is greater by $\frac{1}{4}$ than the difference between the preceding pair. If one happens to try the correct rule straight off, success may come quickly. If on the other hand one begins with rules that are incorrect, success will be delayed and may even be precluded (See Campbell, 1963, 1964). In the survey of the literature pertaining to irrelevant data in problem solving it was suggested that the extent to which irrelevant data will hinder solution will depend, among other things, on whether the irrelevant data are "obviously irrelevant" and on whether they suggest alternative but incorrect solutions. The present study indicates that these two influences can be interdependent: irrelevant data may suggest alternative but incorrect solutions <u>by virtue of</u> their not being "obviously irrelevant". (It may be noted that the irrelevant numbers in these number series correction items are much closer to the "noise" of information theory than are the irrelevant letters of the code items. The number that is odd-man-out - while not "random" - does perturb the series and make its structure less clear.)

Were the irrelevant numbers to be made different in some way from the relevant numbers, much of the difficulty of the number series correction items might disappear. Thus in the item 10 12 14 96 16 18 the 96 can be expected to stand out because of its size, and therefore to be readily recognised as irrelevant. If on the other hand a subject finds a number "perceptually obvious" that happens not to be the irrelevant number, he may find the item excessively difficult. In item (1) above the 8 may attract attention, it being the only even number. It is not the odd-man-out, but the subject may persist in trying to make it so because it is different from the other numbers in this way. These further speculations were never put to experimental test, the number series correction items being abandoned at an early stage due to the difficulty of determining how many numbers must be given for the data to be "just sufficient" for solution, and the fact that a rule can be found, mathematically, to fit any set of numbers.

SAMPLE ITEMS -- NUMBER SERIES CORRECTION

The items in this test are number series correction items, e.g.

Example I:

... ... 12 20 25 30 42 56

You are to find the number in the series that does NOT follow the rule of construction for the series - the number that is the "odd-man-out" - and then put a cross through this number.

You are also to fill in the first two numbers in the series.

Taking Example I, if the number 25 is crossed out then the other numbers fall into a series where the difference between any one number and the one after it is two more than the difference between that number and the one before it:

12		20		30		42		56
	+8		+10		+12		+14	

In accordance with this rule, the first two numbers in the series will be 2 and 6. The number 25 should therefore be crossed out, and the numbers 2 and 6 inserted in the blank spaces.

Now try Example II for yourself.

Example II:

... ... 8 11 22 25 36 50

Overleaf are items (1)-(8). You will be given 15 minutes in which to do as many of them as you can. They are all to be answered in the same way as the examples above. The "odd-man-out" is never the first or last of the numbers given.

DO NOT TURN OVER UNTIL INSTRUCTED

SAMPLE ITEMS - NUMBER SERIES

The items in this test are number series, e.g.

Example I:

... ... 12 20 30 42 56

You are to find the rule for the series and, using this rule, fill in the first two numbers in the series.

Here the difference between any one number and the one after it is always two more than the difference between that number and the one before it:

12		20		30		42		56
	+8		+10		+12		+14	

In accordance with this rule, the first two numbers in the series will be 2 and 6. These numbers should now be written into the blank spaces.

Now try Example II for yourself.

Example II:

... ... 8 11 22 25 50

Overleaf are items (1)- 8). You will be given 10 minutes in which to do as many as you can. They are all to be answered in the same way as the examples above.

DO NOT TURN OVER UNTIL INSTRUCTED