The Role of Memory in Originality

by

James G. Kyle

Thesis submitted as requirements for Ph.D.

Bedford College

University of London

1975

Acknowledgements

I would like to express my thanks to Dr. John Wilding for his patient help throughout the thesis.

I wish also to express my gratitude to the schoolchildren and students who gave their time to the study.

Thanks also to Mrs. V. White who typed the thesis.

Contents | Page

Abstract

The research concerned the application of experimental methods to the study of a cognitive ability. Creativity is normally measured by production and one assumes that information in memory influences the occurrence of original ideas and solutions. Experiment 1 indicated that those performing better on standard measures of creativity had no better gross memory abilities than others - they did not recall usual or unusual words more accurately. The production of unusual words however, was shown to relate to creativity. Experiments 2 and 3 confirmed these findings adding evidence concerning the positive relation of originality and vocabulary. This latter demands access to semantic features in memory.

The final three experiments considered aspects of this problem of access. Experiment 4 examined whether ability to use different cues or codes (visual, acoustic and semantic) was complementary to originality. No relation was found but insight was gained into factors in the memory base whereby one generates more words given a semantic code, though more words are available given a visual code. A model of "distance" was proposed: items are generated according to their distance from a starting point, in conjunction with their activity level or familiarity to the subject. Experiment 5 examined naming latency for recognition of words, given code information. A significant effect of code, semantic faster than acoustic faster than visual, was found, supporting the model, while word frequency, a crude measure of familiarity, was also significant. Experiment 6 showed that the effects were not due to the codes themselves but could be realistically located in the memory base.

The usefulness of the model was discussed and it was applied to the results of the earlier experiments. The final conclusions were that the extent or frequency of the search determines the output and that creative individuals utilise the search process to a greater extent.

INTRODUCTION

The difficulty of explaining the phenomenon of creativity has been with us for a very long time and probably because of our inability to explain it satisfactorily society has accorded it a position of almost mystical reverence. While accepting that creative abilities are of tremendous importance to the growth of a community, it seems only fair to state at the outset that there seems to be no scientific reason or evidence to support the notion that the creative individual stands apart from his fellows in any mysterious way. Why should this be so? In order to state more clearly the approach adopted in this project, let us first examine the methods of approach which have contributed to the current ideas on the subject and attempt to elicit the rationale for this study.

Wallach (1970) suggests a useful classification for characterising the different contributions to the subject. He calls the work done, research on Persons, Products or Processes. "Persons" refers to the question "What is the creative individual like?" - studies which have looked at the personalities of the people judged to be creative. "Products" refers to the question "What is a creative contribution?" and therefore looks at the situations where it may arise. "Processes" asks "What mechanisms are employed by the individual in being creative?" and here we should be looking at the cognitive interior of the person.

The earliest studies mainly involved the "persons" approach (Freud's, 1910, study of Leonardo da Vinci) and the conclusions painted the picture of creativity as a gift allowed only to a select few. In 1926, Wallas published his perennially quoted book, The Art of Thought, and in it he set out an approach to the study of processes which was widely popular until the 1950's upsurge in research. It is mentioned here because it ties in very neatly with the model proposed in the later stages of this research. His outline was based on the analysis of the insights of the great geniuses on their creativity, and particularly on the study of Poincare's mathematical discoveries. What he proposed was that there were 4 stages in the process of creating or solving a problem:

(1) Preparation - in which the problem is investigated in all directions.

(2) Incubation - during which the problem is not consciously thought about.

(3) Illumination - when the idea or solution appears.

(4) Verification - in which the validity of the idea was tested and the idea itself was reduced to exact form.

This stage model was particularly useful as it did agree with the majority of observations of the creative process (see Ghiselin 1952, for a large collection of these observations).

This conceptualisation set the scene for a number of introspective studies of the creative act (e.g. Catherine Patrick 1935 & 1937) but in real terms there was no great progress in our knowledge till the 1950's.

At this point in time, it may be surmised, the definition of creativity was not seen as the most important starting place of a study - it seems to have been more or less tacitly assumed that creativity could be easily distinguished as the manifestation of genius. It was not until the problems of creativity vs intelligence came upon the scene that it was necessary to delineate the bounds of creativity. It is for this reason that a definition has not been dealt with so far.

Before entering upon the explosion in research on creativity in the 1950's, there remains one thread of the story still to be mentioned. This concerns the massive claims made for the intelligence-testing approach to cognitive abilities manifested in the work of such people as Spearman (1927). The most relevant aspect here is seen in the vast longitudinal study set up in 1923 by Terman, which was to run for over 30 years. The study was called "The Genetic Study of Genius" and this conveys the tone of the research. It followed up 1500 gifted children, defined by their high IQ and compared them at various stages to a control group of lower intelligence.

The results were as predicted, to a large extent - the 1500 had more academic and monetary success but strangely, no members of the group had distinguished themselves creatively. Terman's explanation was that:-

> ".... science and scholarship are growing so highly specialised that eminence is becoming more difficult to obtain. Conceivably, if Darwin were living today he might be just another specialist in a restricted field of Biology."

This could not be completely accepted and Shapiro (1968) reflects the feelings that arose:-

> "Since the subjects included many with IQ's far in excess of 140 (the traditional cut-off point for genius) surely one might have expected rare creative attainments from at least a few."

In 1950, Guilford as President of the APA set in motion a very ambitious project to investigate how people think. His address suggested that too much evaluation was based on a criterion of convergent test performance at the expense of divergent thinking, which was probably a more realistic criterion when applied to the real world. His point was not new; Simpson in 1922 had expressed similar fears about intelligence testing. But several factors combined to make the time ripe for Guilford's contribution. Not only did he, over the next 15 years expand his model of the structure of the intellect, but other psychologists and educationalists became interested in the importance of developing the natural resources of creative talent. Hudson (1966) calls it a

"... diffuse cultural ground swell, elevating the scientist from the status of technician to that of cultural hero." (p.120) This interest probably reached a peak in the early 1960's.

To a large extent this line of research looked at the "products" aspect, though as an offshoot an amount of work was done on the "persons" approach. As the clash with intelligence testers drew closer, it now became important to define precisely what creativity was seen to be. A sample of these definitions is now reported in order to make clear the nature of the thing investigated in this research report.

Koestler (1964) says:

> "All creativity is bisociation - a process whereby an item at first belonging in one matrix is suddenly perceived as belonging to a second entirely different matrix."

Mednick (1962) from the point of view of an association theory, defines creativity,

> ".... as involving the formation of associations between stimulus and response which are characterised by the fact that the elements linked together are not normally associated."

Finally, directly relating it to the notion of products, Mackinnon (1962) says:

> ".... 1) it involves a response or idea that is novel or at the least statistically infrequent, 2) it must be to some extent adaptive to or of reality ...
> 3) it involves a sustaining of the original insight, an evaluation and an elaboration of it, developing of it to the full." (p485)

In much of the research that follows the first two points are used as the criterion, the time limit or the nature of the task not allowing the third aspect to contribute to the assessment. In summary, creative thinking involved divergent thinking and the notion of combining two or more items to form a completely new solution.

Guilford's model of the intellect (Guilford and Hoepfner, 1971) allows divergent thinking as one of five general processes. These, defined by factor analytic techniques are:-

Cognition - the perceiving of data from the outside world.

Memory - the retention of what has been perceived.

Convergent thinking - the picking of one (hopefully correct) answer from a number of alternatives.

Divergent thinking - generation of logical alternatives from given information where the emphasis is on variety, quantity and relevance.

Evaluation - reaching decisions about the appropriateness of the outcome of any of the above four.

Although Guilford talks here of processes his research is probably better seen as concerned with products, since the analysis is of the products of various tests and the processes he quotes are simply restatements of the characteristics of the tests.

Further to these 5 classifications, factor analysis allowed him to isolate subclassifications relating to each of them. In the case of Divergent production these were known as: word fluency, associational fluency, ideational fluency, expressional fluency, spontaneous flexibility, adaptive flexibility, redefinition and originality. Though Guilford (1967) has evidence for the existence of all of these, Wallach reviews some research which casts some questions on the usefulness of differentiating these abilities in the context of creativity. The argument centres round the question of whether Guilford's divergent production measures correlate with one another to a greater extent than they do with measures of convergent production, such as intelligence tests.

The type of work that has been done follows the approach of Thorndike (1963) who analysed results obtained by Guilford and Christensen (1956). They found 3 factors in the convergent thinking domain - verbal comprehension, general reasoning and education-verbal correlation - and 5 factors in the divergent domain - word fluency, associational fluency, ideational fluency, expressional fluency and originality. Thorndike took the 2 highest-loading tests as representing each factor and he found that the average correlation among the convergent measures was .43, among the divergent measures was .27 and between the 2 sets of measures was .24. It would seem that the divergent measures have little variance in common, apart from that shared with the convergent measures.

Ward (1966) in an analysis of a study by Guilford et al. (1957), obtained average correlations among the convergent measures of .24, among the divergent of .18 and between the 2 sets of measures, of .15, thus supporting Thorndike's findings. It is therefore an open question as to whether the divergent measures as a whole are something set apart.

However, if we look at the measures individually a more hopeful picture emerges. The 3 fluency factors appearing in the study by Wilson et al. (1954) have an average intercorrelation of .23 while their relation with convergent measures is .14. When only ideational fluency (the ability to generate, within a limited time, ideas that will fulfil particular requirements, such as uses for bricks) was found as in Guilford et al. (1957), the correlation with convergent measures is .01. In contrast, Thorndike finds that the word fluency factor correlates about twice as strongly with the convergent measures as with the divergent ones. As a result, removing the word fluency factor we find that there is a much higher correlation within the other fluency factors than between them and convergent measures (from Guilford and Christensen's study - .31 among fluency and .22 with convergent thinking; from Wilson et al. - .28 among fluency and .08 with convergent thinking). It looks as if discarding the word fluency factor gives a more reliable factor of divergent thinking.

If we follow Wallach's approach as above, there are similar arguments, relating to the factors of flexibility showing that they do not intercorrelate to a greater extent than they do with convergent measures. Similarly with originality (the making of responses which are statistically unique or unusual) there are doubts as to its contribution separate from convergent measures. However a report by Barron (1963) showed that when the effects of intelligence were partialled out from the originality score the cognitive dimension remaining was best definable as ideational fluency. Thus it seems that the tendency to give unusual, original responses is partly a function of greater response fluency and productivity and that this aspect of originality and the fluency measures themselves are distinguishable from convergent thinking.

Wallach (1970) present other evidence of a similar nature which allows him to further analyse the factors Guilford has related to divergent thinking and to reject spontaneous flexibility, adaptive flexibility, redefinition and word fluency. Of the other fluencies he concludes:-

> ".... it is the ideational fluency notion that seems to define the kind of cognitive functioning that is most clearly independent of convergent thinking" (p.1223).

Of originality he concludes:

> "It may be that originality measures will tend to exhibit independence from intelligence if they depend upon assessing the uniqueness or unusualness of the ideas that a person generates whereas they will tend to show relationships with intelligence if they depend on a judgement of cleverness or facility in choice of words in expressing one's ideas" (p.1223).

It may be useful at this point to digress slightly in order to add Guilford's findings concerning word fluency, as it relates to the measures used in this thesis. Guilford and Christensen (1957) and Christensen and Guilford (1963) included an analysis of two types of measure relating to fluency - these were thing listing and word listing. There were 3 conditions of thing listing where the subject had to produce as many words as possible in a certain time according to 3 levels of constraint e.g.

low restriction - name things which are round;
medium " - name things which are round and soft;
high " - name things which are round, soft and white.

Word listing had 4 levels, in terms of the number of prescribed letters to be used in the listing -

O - no letter specified
I - one letter specified e.g. each word should contain S
II - two letters specified e.g. each word should contain E and M
III - three letters specified e.g. each word to contain S, T and B.

For thing listing, they found that the loading on the factor of ideational fluency (IF) was maximal at the moderate level of restriction, the figures being .33, .72 and .55 for low, moderate and high respectively. The extent of the restriction of class caused by the introduction of these restraints is, as they point out, difficult to gauge accurately, and this does not allow an easy conclusion to be drawn.

For word listing, with no restriction there was only very slight variance in factor word fluency (.22) - according to Wallach and others, above, this factor is more related to convergent measures - but with greater restriction loadings of .72, .73 and .65 respectively are obtained. The tests with restrictions had zero loadings on the factor of ideational fluency and there was a loading of .25 when there were no restrictions. Though Guilford takes this as evidence that IF is not involved in any genuine word fluency test, he finds that when the word listing tests are grouped into one score and a targeted solution obtained the loading on IF rises to .36. One further relation should be mentioned: verbal comprehension had a zero loading in the test with no restrictions, .24 in the test with one, .47 with two and .59 with three. This probably confirms the nature of the relation with convergent measures.

In summary, this analysis goes some of the way towards elucidating the nature of fluency in the context of divergent versus convergent production. Accepting Wallach's ideas that word fluency is on the whole more related to convergent production, and this is supported by the final points above concerning the relation with verbal comprehension, it is clear that the increase in the number of restrictions corresponds to a movement towards tapping verbal convergent abilities - this is especially so of word listing. With respect to the factor of IF, the difference between word listing and thing listing is difficult to understand, except in that thing listing may involve semantic cues for word production, while word listing restrictions are symbolic. Overall though there seems to be a degree of relation to ideational fluency, which would appear to be the ability most closely related to creativity.

Returning to the original review, it can be seen that the problem of isolating the factors relating only to creativity has not been adequately solved. There are a number of other approaches which have arisen and which have had some degree of success in answering the relevant questions.

The work of E.P. Torrance is similar in many ways to that of Guilford in that his method of assessing creative abilities is almost the same and the tests themselves in many cases are exactly alike. Where Torrance differs is in the fact of his purpose - his aim is to educate for creativity, thus the slant of his work is to Education and children. He also differs in his characteristic use of an index to describe an individual's performance on the measures used. Some of his materials are verbal, others figural, but as far as possible all are subjected to the same type of scoring - for fluency, flexibility, originality and elaboration. The first two are Guilford's ideational fluency and spontaneous flexibility (the ability to vary one's ideas over a wide range, even though this is not specifically asked for in the test - giving a range of different categories rather than instances of the same category in "Uses for a Brick") and the third is the normal statistical infrequency score for originality. Elaboration is the extent to which the subject fills in all the details of the proposed idea. Examples of the tests to which these scores are applied are: Product Improvement, where the subject has to suggest ways of improving a particular toy to make it more enjoyable to play with, and: Just Suppose, where the subject is presented with an improbable situation (like strings hanging from clouds) and asked to derive the consequences.

The method of calculating the index assumes an equal importance of the four modes above, and simply sums over all the tests in the battery all of the scores for fluency, flexibility etc.

As with the Guilford measures the difficulties with the Torrance (1966) measures lie in their relation to intelligence and convergent measures in general. Though Cicirelli (1965) has shown an average correlation of .37 between the four types of score across verbal and figural tests, this is probably an overestimate because of the nature of the part-whole relationship between fluency and originality; and at the same time, he found the correlations of these indicators with various tests of intelligence, all significant and in the .20's and .30's. Yamamoto (1965) reports correlations of the index score with intelligence of .33 and .39, for a sample of 600. Yamamoto and Chimbidis (1966) studied a sample of nearly 800 fifth-grade school children and when they initially found substantial correlations between intelligence and achievement test score, they calculated partial correlations between creativity index and achievement, holding intelligence constant. The result was that all of the relationships turned out to be near zero, thus illustrating how little the index score was adding to the intelligence test's relation to achievement. This last finding goes against Torrance's "threshold of intelligence" explanation for the sizeable correlations which have been found between creativity and intelligence (it is of the form that above a certain level of intelligence, academic achievement is more strongly related to creativity measures than to further increases in intelligence). It seems strange that Torrance wishes to relate his tests of divergent thinking to convergent measures of school achievement. However, the message is clear - the index does not create a separate entity from convergent thinking.

One way of dealing with these findings from the view of creativity research, is as Wallach (1970) suggests: perhaps the components of the index do not all tap the divergent processes. The conclusion from the studies of Guilford's tests was that only ideational fluency and possibly originality, cohered as a dimension separate from the convergent thinking dimension. The implication is that if elaboration and flexibility are removed from the index, a different picture will emerge.

A study of 150, 13-year old boys conducted by the author, Kyle (1970), used this approach. The creativity measures were selected from Torrance's battery and they were: Figure completion (sets of two or three lines are given and the subject has to complete the picture), Unusual Uses for tin cans, Impossibilities (asking the subject to suggest the things which he could imagine were impossible), Circles (small circles were presented on a page and the subject had to form an object which had a circle as an important part), Consequences (as in the Just Suppose) and finally the Ask-and-Guess task, where a nursery rhyme picture is presented and the subject must:

1. ask as many questions about the picture as possible.
2. state as many possible causes of the events depicted in the scene.
3. give as many consequences as he can, of the action.

Responses were scored for Originality and Fluency and an index for each test based on these two measures was calculated. This consisted of:

$$\text{Index} = \frac{\text{Originality}}{\text{Fluency x max. score per response}} \times 100$$

Average correlation between creativity measures was .27 and average between creativity and Cattell's Culture Fair Intelligence test (which is very highly loaded on 'g') was .04. Thus it looks to be support for Wallach's view:

> "An index score concerned only with ideational fluency and with fluency-related forms of originality would, in this view, exhibit considerably greater orthogonality from intelligence than an index score that is concerned with spontaneous flexibility and with elaboration as well" (p.1233).

The work of Getzels and Jackson (1962), though normally quoted and discussed in a review of the creativity literature, is subject to the same type of criticism that has been levelled at the other studies mentioned so far and also to further more serious criticism concerning their interpretation of the intelligence dimension and their subsequent treatment of groups of subjects arranged by their membership of these groups. The criticisms cast severe doubts on their findings; however they are dealt with adequately by Burt (1962), Cronbach (1962), Wallach and Kogan (1965) and others.

One other important line of research has been that concerned with an associative theory of the creative processes. This derives its inspiration from the work of Maltzman (1960), who proposed a concept called the associative gradient, which determines the tendency for unusual responses. For example, if a subject is asked to produce more unlikely responses in a specific task, then with decreasing commonness of response the gradient of response probability starts lower and falls off more gradually while if the instruction to give more common responses is presented then the gradient starts higher and falls off more steeply. The implication is that subjects have this gradient normally which determines their output.

The idea may be a useful one to describe the type of output that can be obtained but it seems lacking in theoretical background and does not give any insight into the more fundamental processes of word or idea production - in many ways it simply restates the "spew" hypothesis of Underwood and Schulz (1960) (discussed in the second part of the review).

Mednick (1962) has followed in the associative tradition and has developed a Remote Associations Test (RAT) which is related to research creativity. The test consists of items involving the subject in supplying the associative link between three words e.g. rat, blue, cottage - the correct response is cheese. Despite being convergent looking, the solution probably requires the generation of a large number of associates till the correct one is found and thus it is related to productivity and by implication, ideational fluency. The most interesting claim made for the RAT is that it depends for its successful completion on the correct deployment of attention. This takes us away from the simple study of products to a more fundamental view of the creative process. Unfortunately, in the type of research done on this aspect the attentional mechanism seems to be only an intervening variable posed between what would have to be a search mechanism and the output, correct answer.

Maltzman, Belloni and Fishbein (1964) prompted their subjects with a dominant associate of the solution word and found that best performance was obtained when the associate directed the subject to the correct cognitive domain. Mendelsohn and Griswold (1964) showed that high RAT scorers exceeded low scorers in the number of anagrams solved when these words had previously been presented as a focal list (to be attended to) and also when presented as a peripheral list (to be ignored).

What these and other studies illustrate is that there may be a more fundamental conceptualisation of the creative process available; in this case they talk about attention but it looks rather more like the extent of search for creative responses (perhaps like the wide categorisers that Cropley, 1966, talks about - the ability to be aware of all the implications of the task elements). This is seen in the research of Wallach and Kogan (1965) where they found that those high on their associative measures of creativity had broader categorising tendencies, measured on a category width measure adapted from Pettigrew (1958).

At this stage it is possible to draw together the lines of these results to give a picture of where this aspect of research has led. In some respects it can be seen as a circle from the original nature of the creative process as set out by Wallas (1926) through the trials of the mental test movement, where the measures of creativity were discussed, and finally, were clarified in form and scoring as measures on a dimension of importance which has not been taken into account in conventional measures of performance, and then back to the beginnings of the awareness of the need for an experimental examination of the processes underlying not only the general divergent thinking, but also the recently discovered aspects known as ideational fluency and originality. The move towards a more fundamental study of these processes is of extreme importance to the experimental psychologist for he can now attempt to relate this problem to the advances made in recent years in the application of information processing ideas.

Before moving on from the traditional study of creativity, it is necessary to point out that this review has not covered all the strands of the argument and has neglected notably, the work done more recently on the Persons aspect. The main reason for this neglect is that the findings have been on the whole confusing, and on occasions, contradictory. There have been two types of approach:

1. the researcher obtained ratings of eminent men, from other people in the field, and then examined their personality, (Cattell and Drevdahl, 1955, Mackinnon 1962)

2. the researcher looked at the personality differences obtained by dividing a population on the basis of their performance on measures of creativity (Getzels and Jackson, 1962, Wallach and Kogan, 1965).

What has been confusing is that no clear profile has been forthcoming. Myden (1959) has found, using the Rorschach, TAT and other projective techniques

> "... they are inner directed and not easily swayed by outside reactions and opinions... they are sensitive to every nuance or reaction as it may pertain to them...."

Mackinnon (1962) finds that creative individuals are

> "... verbally skillful, interested in communicating with others, and accurate in so doing; intellectually curious and relatively disinterested in policing their own impulses...."

Guilford (1963) says

> "There is some evidence that creative people are more autonomous than others, more self-sufficient, more independent in judgement.... more self-controlled and possibly more emotionally sensitive and more introverted."

It looks very much as if Myden sees the creative individual as introverted, Mackinnon sees him as extroverted and Guilford sees him as introverted again! I think this is the justified basis of dissatisfaction with the results produced by this approach. Perhaps creative people differ for every different field in which they are creative.

To summarise this section, it can be seen that creativity can be meaningfully discussed if we limit claims to something called ideational fluency plus some originality. The emphasis must now be towards an analysis of the actual processes which underpin the type of items created, the method of generation and possibly those underlying the method of storing the elements which now form the creative product.

The progress in the study of Memory

It is now appropriate to look at the advances made in the field of Memory research, in the hope that it will give some indication of the processes which contribute to the storage of items and their final generation as creative responses. The idea underlying a look at creativity in this way comes from the definitions used by researchers in the creativity field (such as Koestler, 1964, when he calls it Bisociation - seeing an item from one context as important in another completely novel context) in the sense that the emphasis is on the use of material already existing in the creative person's mind. Therefore the approach being taken is that the creative individual is someone who is capable of taking items from his memory storage system and applying them in a manner and in situations, where the contribution of this application is novel and helpful to the task in hand. The relevant questions then are whether the creative person somehow stores unusual combinations of items in the first place, or whether his ability is to gain access to uncommon items in his memory or whether his ability to evaluate the importance or relevance of the items which can be obtained from memory, is somehow more developed than other individuals'.

The particular stimulus for this research derives from studies such as that of De Groot (1965) in which he examined the attributes which went to make a chess master different from an experienced chess player. One of his most interesting findings concerned a test of memory when he presented subjects of the "master" class and subjects of the experienced class, with positions from actual chess matches for varying intervals of time from 2 to 15 seconds, and then after allowing 30 seconds for organisation, he asked for recall of the exact positions of all the pieces previously shown. The recall of the masters was markedly superior to that of the experienced players, but also their method of recall was different. They seemed to have a more dynamic grasp of the situation and could relate it to their experience more easily than the experienced players - whether this was due to their initial perception of the stimulus or their later organisation, in recall, of it, is not clear. They tended to make general observations of the run of play which then allowed them to deal with smaller sections of the situation - rather in the way of higher-order memory units.

This tends to tie in with the biographical studies where it is common to find that the creative person's account involves grouping his perception in a particular unusual way, which involved a more dynamic use of the cues in the task situation and of the ideas being produced from his memory.

It is therefore important to examine the results of research in the processes involved generally in memory, in order to provide a context for this theoretical approach.

Studies of Memory Processes

One of the greatest difficulties in a review of this nature is that research in this field has become so popular that there has been a proliferation of findings based on a large number of methodological approaches, and the advances made in this area are clouded by the very diversity of the theory. Inevitably then a venture of this kind must stop short of a complete charting of the development of research on memory and concentrate on the aspects of direct relevance to the undertaking here. This means dealing with the Memory Base (more or less very long term memory, representing the items which the individual has known for a good deal of time, reflecting his experiences and skills which he brings to any task situation), and the aspects of Storage and Retrieval which are then logically of such importance.

Three main theoretical approaches will be dealt with: Learning theories, Association Networks and Information Processing - the first two derive from the same Psychological background and complement one another, while the third represents the approach adopted by the great majority of the research workers in the field.

LEARNING THEORIES

Historically, the roots of research on verbal learning and memory are to be found in Associationism. The first experiments by Ebbinghaus on rote learning emphasised the acquisition and retention of serial associations. With the development of behaviourism, the study of verbal learning became identified with the study of stimulus-response associationism. What this approach gave to this study was a methodology of laboratory experimentation and a descriptive framework for the phenomena of learning and retention and for the subject's performance therein. Some measure of the importance of this approach, which has persisted to the present-day, can be gained from reviews such as McGeoch and Irion (1952)*.

One model which illustrates this approach albeit in a more extreme form than it is usually encountered today, is the work of Rosenblatt (1958) and his Perceptron model. As the name suggests it is a model of perception but in terms of memory storage its principles are equally applicable to verbal material. What the model assumes is that patterns or representations of the stimulus do not exist per se in the nervous system. The activity intervening between the presentation of a pattern and the response to the pattern is seen as a transference of information from the sensory to the appropriate motor system, through pathways selected by reinforcement.

*McGeoch, J.A. and Irion A.L. The Psychology of Human Learning, New York, Longmans, 1952.

Thus according to the traditional learning theory approach, memory is the evocation of a correct response, by the appropriate transference of energy from sensory to motor components of the nervous system. Though this outline is extreme in some respects, when the intervening activity is given some conceptual substance, it can be an extremely important theory of memory (Pribram, 1971 to be discussed later).

Looking further into the principles of the Perceptron, its relation to Associationist tradition becomes clear. The development of such a memory system is founded on certain basic axioms (p.388-9).

1. "... at birth the construction of the most important networks is largely random, subject to a minimum of genetic constraints."
2. "... after a period of neural activity the probability that a stimulus applied to one set of cells will cause a response in another set is likely to change..."
3. "Through exposure to a large sample of stimuli those which are most 'similar'... will tend to form pathways to the same set of responding cells. Those which are 'dissimilar' will... develop connexions to different sets of cells".
4. "The application of positive and/or negative reinforcement... may facilitate or hinder whatever formation of connexions is currently in progress".
5. To allow for stimulus equivalence e.g. in recognition of different forms of handwriting, he adds

> "... similarity... is represented by a tendency of stimuli to activate the same set of cells."
> Therefore it is not a necessary attribute of the physical form of the stimulus. A stimulus such as 'q' is related more to 'Q' than to 'g'.

The model explains the development of a memory system in operational terms for each individual, in that from random beginnings the relation of items in memory is set up by their pattern of usage. In many ways this is supportive of Tulving's (1964) proposals of Subjective Organisation, to account for the Subject's order of recall. On the access side the model is weak, and less helpful in a search for an explanation of the generation of ideas, since it tends to suggest that in order to output a particular item we must have available the correct half of the stimulus-response relation. This seems to go completely against the notion of bisociation. In summary, what this model contributes is the notion that memory is organised in terms of the relations which have been <u>functionally important</u> in the experience of the individual and also an indication that it may not be necessary to think in concrete terms when we consider what exactly a memory consists of.

The more conventional use of learning theory principles in this field can be detected in the work of McGeoch, Postman and Underwood, though the work of Underwood is more directly related to the purpose here. The emphasis here has been to find which variables influence retention and account for the phenomena of organisation found in recall. One of the variables that has attracted a great deal of attention is meaningfulness, the extent to which the subject finds the stimulus "meaningful".

Glaze (1928) collected norms for the association value of a large number of nonsense syllables, defining meaningfulness (M) as the extent to which subjects could produce an association to the syllable in a fixed interval of time. Noble (1952) introduced a slightly different method of estimating M - the production method - whereby subjects had to produce as many associations to the stimulus word as possible in 60 secs., the number produced indicating its M-value. His stimuli were words drawn from high and low frequencies in the Thorndike-Lorge word count (1944) and also paralogs (like tarop, gokem etc.). Other methods have been used, notably by Hull (1933) and Noble, Stockwell and Pryer (1957), but in all cases the correlations among the values obtained is high (Noble et al., 1957). As an experimental variable the findings have been very clear cut: ease of learning and recall seems to be a direct function of M (McGeoch, 1930, Dowling and Braun, 1957, for serial learning, Noble and McNeely, 1957, for paired associates).

Underwood and Schulz (1960) however, suggest that there is a more fundamental variable at work here, namely word frequency. Evidence supporting word frequency as a variable affecting learning and recall, is provided by Hall (1954), Bousfield and Cohen (1955) and Bousfield, Cohen and Whitmarsh (1958). All these studies show recall as a function of

Thorndike-Lorge word frequency, with high frequencies being significantly better recalled. Correlations of word frequency and M show a clear relationship of high frequency - high M ... (Underwood, 1959, Cofer and Shevitz, 1952). Thus there is strong support that word frequency, or frequency of experience of the subject is the fundamental variable relating the different notions of M, and their effect on learning.

These considerations led Underwood and Schulz to propose the Spew hypothesis:

> "the order of availability of verbal units is directly related to the frequency with which the units have been experienced". (p.86)

Data supporting this notion is provided by Johnson (1956) and Howes (1957). Both examined the relationship between the position of a word in the Kent-Rosanoff response lists of association and its Thorndike-Lorge frequency. Both had similar findings and Howes estimates the correlation as .94. Bousfield and Barclay (1950) in a study where subjects had to name as many members of a class as possible e.g. birds, show that the actual order of emission and frequency are highly related.

This Spew hypothesis seems to express the same type of findings as Mednick (1962) has obtained with creative associations, in that the more unusual responses tended to be produced towards the end of the subject's response list. Perhaps the frequency of experience determines the accessibility of the response.

Underwood and Schulz go on to develop the hypothesis in relation to a two stage model of verbal learning, these being response learning and association. The assumption that all learning can be explained with these two stages raises a large number of complex questions, which would take this review far from its theme. At present, it is more relevant to examine further findings of the learning approach which relate to the organisation and retrieval of items in memory.

Much of the other work done in this vein, tends only to add to the demonstration of the effects on memory of pre-experimental habits - associative, linguistic and conceptual - which the subject uses. The use of free recall learning allows the experimenter to study these habits and also to note the types of organisation employed in retrieval. Deese (1959) reports a study of inter-item associative strength - "... the average relative frequency with which all items in a list tend to elicit all other items in the same list as free associates." (p.305). Having controlled for word frequency, he finds that there is a positive high correlation (.88) between inter-item associative strength and the number of words recalled per list, and a negative correlation (-.48) between inter-item associative strength and number of extra-list intrusions in recall. Though it may be argued that the common associates act as a mnemonic and therefore increase recall, rather than the more mystical associative strength, the demonstration of the subjective organisation remains important.

A number of studies have illustrated the phenomena of clustering in recall (these are reviewed at length in Cofer, 1967 and Postman, 1964). Briefly, the approach has been to present the subject with a list of words belonging to various culturally defined categories, placed randomly throughout the list, and ask for free recall. Typically what is found is that the subject groups his responses along the lines of the categories rather than on input ordering. The degree of clustering is the extent to which the amount of categorical groupings exceeds the level to be expected by chance. Bousfield, Cohen and Whitmarsh (1958) show that the greater the associational value of the words to the category, the greater will be the degree of clustering. Again this can be interpreted in a different way (higher-order memory units, Miller 1956) and it begins to point to the limits of Learning theory approaches to memory.

In general, this type of approach has most to contribute when expressed as interference theory (McGeoch 1932), in order to study forgetting. Based on the finding that the learning of laboratory materials tends to interfere with memory for other laboratory material that was learned before and after it, it has built on the easy interpretation of the associative framework. However this has led to the great weaknesses in their view of the retrieval process and there is no explanation of interference, deduction and reconstruction in the acts of remembering. What can be gained from the approach as has been pointed out above, is an explanation of the development of a memory system and the mechanisms whereby, given an open-ended task, items are produced from storage, in particular orders (the Spew hypothesis).

ASSOCIATION NETWORKS

Though the idea of having some form of associative network as a central unit in memory has been proposed, it is only with the development of large scale storing facilities, offered by computers, that the full implications of such a system have been appreciated. The basic stimulus to this development was the appearance of LISP as a list-processing computer language, in the early 1960's. What was so useful about LISP, was that its basic structure was an ordered pair whereby two memory locations may be conjoined. Also, standard to LISP is the property list which relates elements by means of particular relations. Thus fundamental to the language are two structures which are extremely similar to unlabeled associations (ordered pairs) and labeled associations (property lists) - labeled associations carry the relation of the two elements e.g actor to object, or actor to location etc. This type of system is used in a very large number of computer models. However, it is only fair to point out as Bower (1972) does:

> "...when a cognitive simulator decides to use a network of associations as his data base, the considerations that are motivating him may be very different from the traditional psychological concerns. That is, cognitive simulators often are not trying to capture the associative trains of thought, ... or other more or less immediate data. Rather they are using an associative network because such representations are considered in computer

science to be desirable for constructing large general purpose data bases."

Be this as it may, the fact that we can specify such associative systems allows an estimation of the feasibility of some of the ideas expressed in the previous section. One such model is that of Kiss (1969).

Though it is obvious from his later contributions (Kiss, 1972) that he wishes to fit his model into the contemporary framework of encoding systems and feature analysers, the initial statement and tests of the associative store are of interest here.

As a preface to this model, it must be pointed out that Kiss bases many of his arguments in the realm of physiological fact. Despite the claims of some psychologists that research and theory can exist on a psychological level without recourse to physiology, he feels there is no virtue in this approach. Indeed the nature of the physiological mechanisms can evaluate the contribution of various models. Though this is a tender topic in this area of psychology, where much of information processing theory exists almost despite physiology, it seems quite possible that the closeness of the "physiological fit" of a model of the data base, may in the end decide which is to be accepted as "correct" for science.

Collins and Quillian (1972) highlight the danger:

> "Computers' interest in psychology stems from their desire to copy the way people think. They analyze people in terms of the way they themselves, at present, work: that is, in terms of strategies, routines and subroutines, pointers, lists etc. In these terms, they feel they can imitate any kind of process from a chemical reaction to a confrontation between ego and id. It is not so important to them whether this is the best way to try to describe how different processes work, because it is the only kind of description that is of use to them" (p.310).

A warning indeed.

Kiss' information retrieval system is governed by stochastic processes, which is intuitively acceptable since there are times when words cannot be retrieved though they are familiar to the subject. His interest is in words and the store contains representations of the words. The exact nature of this representation is not specified, though it is not like a computer address.

> "Human memory is organized on the basis of its contents rather than on the basis of the location of the contents" Kiss, 1972 (p.329).

And, as if to emphasise the link with learning approaches -

> "The problem of accessing and retrieving information is best represented as a function which, when applied to the given information as an argument, yields the required information" (p.329).

Each word in the store has a certain level of activity, which changes over time and the state of the word store is specified, at any instance in time, by the current level of activity of all the words in it. The word store operates by going from state to state (transitions) as the level of activity of the words changes, due to certain influences. These may be (a) information transmitted from the sense organs affecting areas of the store, (b) higher level functioning, presumably self-initiated - the thought processes, (c) intrinsic influences between words in the store. This third influence is probably the most difficult to imagine and determine, though it is the simplest mathematically, and in fact this is the area upon which Kiss concentrates.

These "free" transitions are stochastic in nature, and therefore, once the initial state of the system has been specified, it is possible to determine the evolution of the system through time. At any moment in this process a word can enter "consciousness", according to its level of activity - the higher the activity level the higher the probability of its selection. This is almost a theory of response strength where probability of selection is a ratio of response strengths.

These ideas have instant appeal as the brain is a constantly active system, and there is evidence of random processes in the form of random threshold fluctuations, internal noise and irregular spontaneous activity (Fatt and Katz, 1952, [redacted], [redacted]). Each word in the store then, is represented by a set of neurons whose behaviour determines the level of activity and whose synaptic interactions determine the transmittance or associative strength of particular words.

Mathematically, he uses a vector Markov process, described by Harris (1963) which, if the assumption is made that state X_{n+1} depends only on X_n *, generates an expectation process which determines the relative proportions of the activities of the words at the asymptote.

In order to evaluate his model in psychological terms, Kiss turns to the traditional associative word structure as detailed by such researchers as Palermo and Jenkins (1964). These norms for large numbers of individuals constitute the degree of concordance in verbal habits, and traditionally for psychologists, by implication the response hierarchy or structure for the individual.

The model conceives of the process of word association in the following way: the stimulus word produces a specific starting state, by initiating activity in certain areas of the network; the system evolves through time until a decision to respond is made; the output is determined by the relative level of activity of that word. He adds:

> "Since this decision is a stochastic one, variability of behaviour is a characteristic feature of the model, in accord with our experiences with verbal behaviour."

*This assumption may be difficult to uphold in certain circumstances, e.g. a word produced at the beginning of an associative listing, of high natural activity, does not tend to be produced again towards the end.

This adds an extremely important sophistication to the Spew hypothesis, in that it could under a stochastic system, explain why subjects do not always repeat the same order of associations. One of the important contributions of the Spew notion was that since it expressed frequency of experience, it could give a picture of the construction of a network of this nature.

Despite the clear implications for individual, Kiss disappointingly looks at group associative structure. In fact, his most recent work (1973) concerns group patterns or the associative thesaurus. His test of the model compared the probabilities of association calculated from the model based on the value matrix (the frequency of particular inter-associations) and the transmittances obtained to the level of three associative links, with the word association norms of Palermo and Jenkins (1964) and Russell and Jenkins (1954). To obtain the value matrix, Kiss had subjects associate to the high and low frequency associates of 4 key words. The comparison of the model's transmittances were then with the norms' relative frequency of association to the 4 key words. Correlations are high (.578 and .669) and clear support is given for the model, by this type of data.

The value of the model has yet to be completely estimated, and the problem of doing this lies in the examination of the individual and only secondarily in the prediction of group norms of association. The relatively simple version of the Spew hypothesis, in terms of frequency of word experience can account for word association norms and the order of production (Howes, 1957, reports a correlation of .94). Therefore a more crucial test of its concurrence with individual behaviour is required for the model, to be completely accepted. From the point of view of physiology, the model is extremely attractive and it calls to mind Estes Stimulus Sampling theory (1960) in that it would provide a similar probabilistic learning model as the particular representative sets of elements come to be attached to particular words. This is similar to Kiss in the states of the associative net which become representative of a particular learning situation and response. Recurrence of the state as in S.S.T., determines the response.

One may note in rounding off the look at association networks, that though few others have contributed to the sophistication of the actual network, there have been interesting suggestions made. Anderson's simulation model of free recall FRAN (1972) based on all-or-none pathways through the network searches along particular chains. This allows an explanation of the characteristic bursts of responses in association tasks, found by Pollio (1968) and others, and the resulting inter-response times accruing. The items within a burst tend to be those which the subject has grouped together in a "subjective unit".

The pauses between successive recall bursts grow progressively longer as though subjects were exhausting the accessible units in memory. Fran searches along a chain of tagged items till it exhausts that subjective unit, it then returns to the starting point to follow another chain and thus produces the longer inter-response times after the bursts. Though the all-or-none nature of the model's pathways may be disputed the notion of the subjective grouping and the corresponding search along a particular line of tagged items, is a powerful one. One in fact, which Kiss (1972) rejects in the form of a random walk in his model.

In summary, the contribution of computer systems to the associationist approach has been an interesting one and does add to the picture of generation of items from memory, as an associative search which can account for many of the findings of the earlier approaches to verbal learning.

INFORMATION PROCESSING

Despite the large amount of speculation and thought devoted to the study of memory, in years gone by, (Paivio, 1971 for a comprehensive review) relatively little progress has been made until recently. Only within the last two decades has the study of memory produced viable explanations; much of this has been due to the development of the information-processing approach. Basically this approach sees the human as a processor of information, rather than a simple stimulus-response mechanism, and though the internal processes of the individual are very much a "black box" of less importance to the cognitive psychologist, they are characterised by functional relationships verified by a number of experimental results. In many respects it is a rejection of the reductionist viewpoint and an attempt to construct a model on one level of scientific knowledge which will account for the results of psychological investigation.

The initial approach as set out by Shannon (1948) with its proposal of information measured in terms of uncertainty (bits) proved to be less useful than at first thought, and only with the more recent attempts to characterise information in terms of chunks, features, semantic markers etc., (as descriptive of the subjective conceptual mechanisms, compare with Tulving's subjective organisation, 1962, previously mentioned) has it been possible to realise more successful models of memory.

Even within the perspective of so short a period of development certain stages of progress can be detected. Though these are not completely separable in temporal respects, they constitute scientific stages of the initial broadening of the concepts of information processing through experiment, the relatively successful attempts to incorporate the findings in viable models of memory, a further period of expansive research, more or less up to the present time, and the first tentative steps to much more accurate models, taking place at the present time.

EARLY RESEARCH

(i) Initially the investigations concerned the capacity of memory and from these arose the notion of how different memory stages could exist within the overall storage-recall idea. Miller (1956) was the first to bring together evidence to support his idea of fixed capacity, of the memory system which in terms of the information theory measure was found to be around 2-3 bits. The method of overcoming this large limitation on the span of immediate memory was to recode the information into higher order units which he called 'chunks'.

More concrete information on the nature of the memory processes came with the distinctions between sensory, short term and long term memory. Sperling (1960) showed that there was some form of visual representation of briefly presented matrices of letters which significantly aided recall of these letters. The storage takes the form of a rapidly decaying "trace" whose items are being lost even in the short period of time (a matter of seconds) during which its items are being recalled. Peterson and Peterson (1959) in a study or recall of consonants without rehearsal illustrated the temporal decay of short term memory storage, (recall drops from 80% after 3 seconds to 15% after 15 seconds.)* Essentially they noted the importance of rehearsal if items are to be retained in memory. Work by Murdock (1962) on the serial recall curve showed that meaningful distinctions could be made between this short term process and a more stable long term store. He presented lists of words of varying length to subjects and asked for recall; the finding was that there was better recall for the first 2 or 3 items in the list and for the last four or five. When recall is delayed by using the technique of Peterson and Peterson (1959), the recency effect (better recall for the last few items) disappears, thus supporting the notion that retention of these items is due to the same process as invoked in the Peterson study, i,e. short term memory. The better recall of the first items is then attributed to the action of long term memory. (also Glanzer and Cunitz, 1966).

*Subsequently there have been some doubts about this simplification - proactive inhibition may be the main factor in the decay.

Despite some opposition to the distinction of these type of stages in memory (Melton, 1963) support has come from the study of brain damaged patients. Milner (1967) reported the case of a patient with damage in the hippocampal area of the brain, who could not store new items in long term memory. While the information was in his immediate memory he could perform normally, but as soon as his attention was distracted and the STM cleared he lost all memory of the task. Shallice and Warrington (1970) report a case of grossly impaired short term memory (STM) but normal long term memory (LTM), thus instead of a normal recency effect in free recall of four or five items he had only better recall on the very last item. Interestingly enough, though this supports the distinction of the different memory systems, it casts some doubt on them as stages of the same process, since it had been previously assumed that one had to use STM in order to put items into LTM.

Looking more closely at these systems of storage certain characteristics have been associated to them as the outcome of research. Though the main line of research on sensory memory has concentrated on visual memory, there is also some evidence for an auditory store of a similar kind storing items for a brief interval. (Broadbent, 1958, Morton, 1970). The distinction is made between iconic (visual) and echoic (auditory) memories. The general characteristics are the same - the storage is devoted to physical aspects of the stimulus. In Sperling's experiments subjects reported items on the basis of their spatial location, Clark (1969) used colours and other experiments have shown that size and shape can be used (Turvey and Kravetz, 1970, Von Wright, 1968). But the limits of the process can be seen when the items have to be processed further in order to be reported, e.g. if asked to report letters or digits in a mixed array, subjects do no better than when asked to report simply as many items as possible from the array. It would seem that this sensory system is confined to analysis of the physical attributes of the stimulus. That is, the encoding is of the physical details.

In STM the type of encoding seems to be acoustic or articulatory. Sperling (1963) found that recall is based on a verbal recoding of the selected contents of the visual information store, which is remembered in the auditory information store and is subject to rehearsal. Conrad (1964) maintained that the errors occurring in recall from STM are systematic, and that items which are acoustically similar are most likely to be confused with one another. Conrad (1964) showed this by illustrating a highly significant correlation between subjective confusions in recognition of auditorily presented letters against background noise, and recall of the same letters from short term memory.

There is evidence that LTM is based on semantic encodings. Baddeley (1966) (and Baddeley and Dale, 1966) conducted a number of studies comparing the effects of acoustic similarity and semantic similarity on STM and on LTM. The stimuli were high frequency words chosen for their similarities. In short term recall acoustically similar words are less well recalled than different words (with a slight effect of semantically similar words), while in long term recall semantically similar words are less well recalled (with no effect of acoustically similar words). Similar results were found by Kintsch and Buschke (1969). The message seems to be clear: there are three systems which seem to correspond to different encoding procedures - sensory, (physical) STM (acoustic) and LTM (semantic). Perhaps the question to ask is whether these encoding procedures are the temporal aspect of the memory system, giving rise to this temporal distinction between the Sensory, STM and LTM. Indeed Postman (1975) investigates this possibility and suggests a "principle of time-dependent encoding: it takes longer to process an item semantically than to process it phonemically." p.300

One further aspect of the research on memory at this stage was the investigation of the nature of the organisation. Although Bousfield and Cohen (1953) showed a correlation between clustering and recall, it did not indicate a causative relation. What is required is to show that the actual procedure of organising affects recall. Tulving's measure of subjective organisation sets out to do this.

The measure of subjective organisation (SO) is a measure of the extent to which subjects remember words in the same order from trial to trial. Tulving (1962) showed that the degree of SO increased over trials in parallel to the degree of recall, and also is significantly greater than the SO that would be expected by chance. Taking into account that the measure of SO is probably an underestimate of the degree of organisation (since it is based on order of items and not simply contiguity), it is quite clear support for the notion that subjects use organisation in recall.

Mandler and Pearlstone (1966) further support this. They showed that subjects yoked to controls and asked to sort cards until their sorting was exactly the same as the control, recalled the same number of words from the cards as the controls despite having about twice as many trials to sort the cards correctly. The implication is that neither time of exposure to the stimuli nor freedom to categorize the words in one's own way is as important as actually categorizing the words, for correct recall. A correlation of .95 was found between total recall and the number of categories recalled, suggesting the use of the categories to retrieve the words.

Tulving (1966) casts light on the extent of the effect of categorizing. Two groups of subjects learned a nine-word list to criterion and then learned a list of 18 words. For one group all the words were new, for the other group the list contained the nine already learned words plus nine new words. Initially the group with the old words had better recall but after 4 trials the other group had caught up and after the eighth trial showed consistently better recall. The explanation suggested by Tulving is that the organisation imposed on the original nine words was not optimal for the list of 18 and the subject's inability to give up these higher order units interfered with his attempts to learn the nine new words. The general conclusion drawn from these studies is that subjects do tend to organise their input into higher order units and that the problem in recall becomes one of retrieving the appropriate category names. Other supportive evidence is provided in the work of Cohen (1966), Tulving and Pearlstone (1966) and Bower et al. (1969) where the category structure was supplied by the experimenter.

EARLY MODELS

(ii) The first attempts at model building tended to restate the findings on the different memory stages and models such as Sperling (1963) and Waugh and Norman (1965) are of this type. The main contributions tended to be of a mathematical nature in the name of rigour and predictive power. (Norman 1970, devotes a whole book to contributions of this type). One particularly influential theoretical model was that of Atkinson and Shiffrin (1965, 1968). Their model consisted of three stages: the sensory buffer (which receives information from the senses in proportion to the exposure time of the stimulus), the memory buffer (a rehearsal scheme where attention is recycled among the items held there) and the long term store (which consists of items copied from the short term memory buffer). Items are lost from the sensory buffer in the absence of rehearsal or special attention. Most of the postulates are directed at the buffer or STS about which most was known at the time and the properties are that it is of constant size, so that as each new item enters an older item must be lost with a certain probability, and that it acts as a push-down store so that older items have greater probability of being lost. At any point in time an item has a fixed probability of being copied into long term store (LTS) and important here is the idea that the item can be in both STS and LTS simultaneously. Mathematically, by the estimation of the various parameters of δ (the probability of losing the item in position 1 of the buffer), θ (the probability of copying an item into LTS) and r (the size of the buffer) a good fit can be obtained for most of the free recall data supplied by Murdock (1962). Three parameters are all that are required to give a good prediction of long term effects also. The advantage in this approach is that it encouraged rigour in the specification of the processes in memory and in return allowed a clear prediction of the raw data of recall. Its weakness lies in the fact of its insensitivity to the organizational principles of the memory system.

Another model which is indicative of this attempt to explain the findings on memory in terms of the structures involved is that of Morton (1970). One or two new ideas were introduced in this model and they are important to this review. The model is based on a model set up to describe word recognition performance (Morton, 1964, 1969, Morton and Broadbent, 1967) - it is therefore a model which deals more with the input and its analysis, rather than the memory base or the acquisition of the items therein. Figure 1 shows the outline of the model.

The basic assumption is that when a verbal response is available as a result of some stimulus, the same final unit operates to produce the response regardless of the source of information which led up to the response. This final unit is termed a Logogen, and it is defined by its output, represented as the set of phonological features and the attributes of visual, acoustic and semantic encodings.

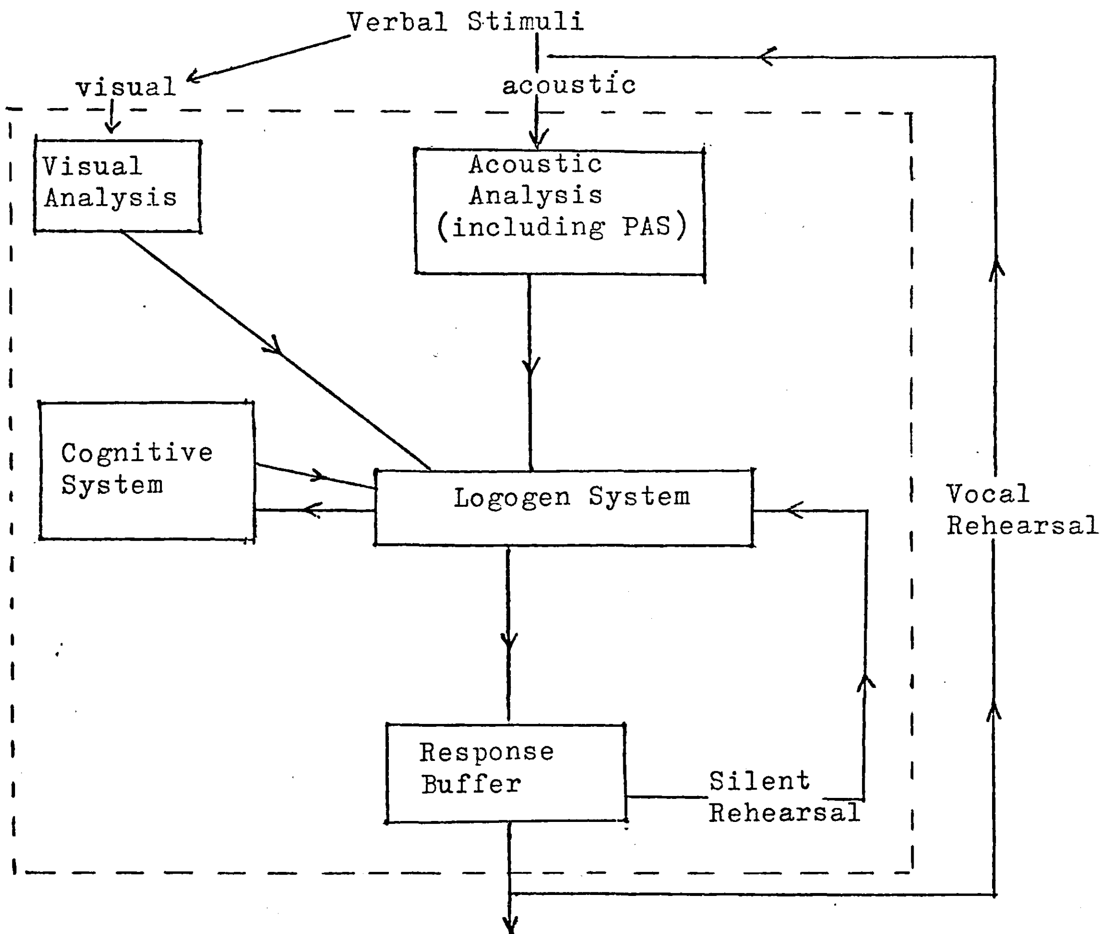


Figure 1. Morton's Logogen Model (Morton, 1970)

(This is similar to the approach adopted by Norman and Rumelhart, 1970, where encoding is by feature extraction and the memory is represented as a system of vectors of the attributes extracted.) The logogen is a counting device which is incremented whenever there is an input of an attribute belonging to any one set. When the count exceeds a certain value a response is made corresponding to the Logogen's response set. Reflecting his work on word recognition there are differential thresholds according to the frequency of use of the Logogen; thus high frequency words have logogens with lower thresholds. This is quite a useful innovation since the problem of word recognition becomes for the subject one of signal detection and this area of psychology has evolved decision models which allow description and prediction of a subject's behaviour. Its incorporation is not unique to Morton's model, indeed Pollack (1959) used decision processes in his study of recognition memory for tones and Murdock (1965) and Norman and Wickelgren (1965) also made contributions along these lines. One important feature of such an addition is that it separates the process of extracting information from a signal, from the process of making a response based on the analysis. This is greatly removed from the theories of Rosenblatt (1958) and the Learning approach in general.

This arrangement will then account for the errors made by a subject in a word recognition task e.g. since there is an overlap between the visual attributes of cat, V_{cat} and sot V_{sot} one would expect that when the stimulus sot is presented in an impoverished fashion the response "cat" will often appear, but not vice versa - the reason being that the Logogen for cat has a lower threshold. The type of analysis suggested to make prediction with this model more accurate is that of Response Strength Analysis (Luce, 1959) whose basic idea is that the probability of a particular response is given by the ratio of the response strength of that particular item to the sum of all the response strengths.

Much of Morton's development of the model is in the realms of STM and concerns the distinction between the precategorical acoustic storage (PAS) and the response buffer, as contributors to the results of studies of short term recall, and is not directly relevant here. However, his argument that visual information from the visual analysis is not immediately transformed into auditory information in some STM system such as proposed by Sperling (1967) is important. It does find immediate support in such a phenomenon as perceptual defence, where meaning is extracted before the appropriate response is made available to the output system. It can be argued that if the response attributes are available for an internal rehearsal of an acoustic nature, then they must be available for output. In this case it is unlikely that there is <u>necessarily</u> direct access from the visual analysers to an auditory store. The model in itself, however does not separate articulatory and acoustic encoding - as the mechanisms of response buffer producing vocal rehearsal and then acoustic analysis, could quite easily be classified as one process of auditory buffering, with access to the LTS (which is in fact, the system used by Atkinson and Shiffrin, 1967, though their emphasis is not on particular aspects of the coding). The data used by Morton (Crowder and Morton, 1969 and Morton and Holloway, 1970) to support his stance in this is unfortunate in that it concerns lists of visually presented digits to be recalled which are unlikely to require any auditory processing; (the reason being that digits are such well-used material that it is quite conceivable that a visual code alone - the number shape - would be quite sufficient to trigger the semantic relevance of the number. The necessity for auditory encoding is greatly diminished) so that they could be quite easily detached from the auditory input of a suffix. Also the basic argument that there can be no explanation for the advantage of auditory presentation over visual, seems to misinterpret or underestimate the extent of the information difference for the subject between visual and auditory stimuli. To the extent that articulatory and acoustic encodings are the initial material used by the individual in learning to communicate, and then the most common types of information encountered by the individual in everyday life, there would seem to be every reason for a better performance on auditory tasks and if this is studied over short periods in memory the effect is likely to be most clearly seen in the recency part of the output list, as the encoded attributes are more accurately preserved in the auditory store.

What Morton has shown is that it is not necessary to assume that all items proceeding from the visual analysis are immediately reprocessed as auditory encodings before there is any access to LTS and its semantic information.

The final part of the system, the Cognitive system, is identified with the LTS in many other models; the information residing there is coded semantically, in accordance with the findings already quoted, though there will be some unspecified degree of sensory coding. The phenomena of clustering, semantic confusability, idiosyncratic recall methods, are all attributed to the cognitive system, but Morton makes no attempt to specify their nature in this storage.

In summary, the Logogen model illustrates the approach taken in response to the initial research findings - the emphasis is on short term memory and the system can be easily represented as a through-process flow diagram. One of the most important points about the model has been discussed above (the necessity of direct access from visual to acoustic analysers). Other aspects which should be noted at this stage are the implicit assumptions that the input stimulus is dealt with in terms of the sensory analysers, as distinct from some form of filter mechanism, and that these encodings are the ones used in the retrieval of specific items, thus broadly explaining the different types of "confusions" found in recall tests. The concept of the Logogen is itself of some importance - it can be compared in many respects to the "name" code suggested later by Posner and others (1973). It is a device for transforming incoming information on the basis of encoded attributes and as such allows no interaction among Logogens. The main characteristic is that it counts up to a certain predetermined threshold value for each Logogen, and outputs the response set at the point when the attribute count threshold is exceeded - the problem of output can be characterised as one determined by a decision process, where a response occurs with a probability based on attribute count values.

One difficulty with the model would seem to be that the role of The Cognitive System in short term recall would predict that the extent of semantic confusability would be as great as in long term recall, which according to Baddeley (1966) is not the case.

It is clear that Morton's model does not seek to explain the findings on organisation in memory except in the sense of encodings and this was generally true of most of the models proposed at that time. However, a model proposed by Kintsch (1970b) goes at least some of the way towards a working conceptualisation of the organisation in free recall and the phenomenon of clustering. His model exists within the framework of Atkinson and Shiffrin's (1965) model of the memory processes, with the difference that he proposes to specify more accurately what constitutes "transfer of information from short term to long term store". The representation of a stimulus item, like a word, is acoustic (following Sperling, 1967) - the particular matrix elicited by this makes accessible markers corresponding to the sensory features and the semantic and syntactic features. The markers associated with the word are then time-tagged. The assumption is that each marker has a familiarity value and that tagging a marker amounts to incrementing this familiarity value. Familiarity values decay exponentially over time and therefore consolidation has an important role in the system.

In retrieval, after the production of the items in STS, given a starting point, a word is chosen at random from those available and its markers scanned. Marker X with the highest familiarity value, is selected and the system moves to an examination of the entry appropriate to that marker. If the markers of X are above a criterion, X is produced as a response. The search continues with the already produced word as its new starting point.

This model shows how LTM may be used to learn to recall a list of words, since pre-existing relations among words guide the search through memory. Items related to one another will have overlapping markers and retrieval of one item would tend to produce the related word in contiguity. The model is not completely testable largely because of the search process postulated, but Kintsch does provide some evidence to support the organizational nature of the system.

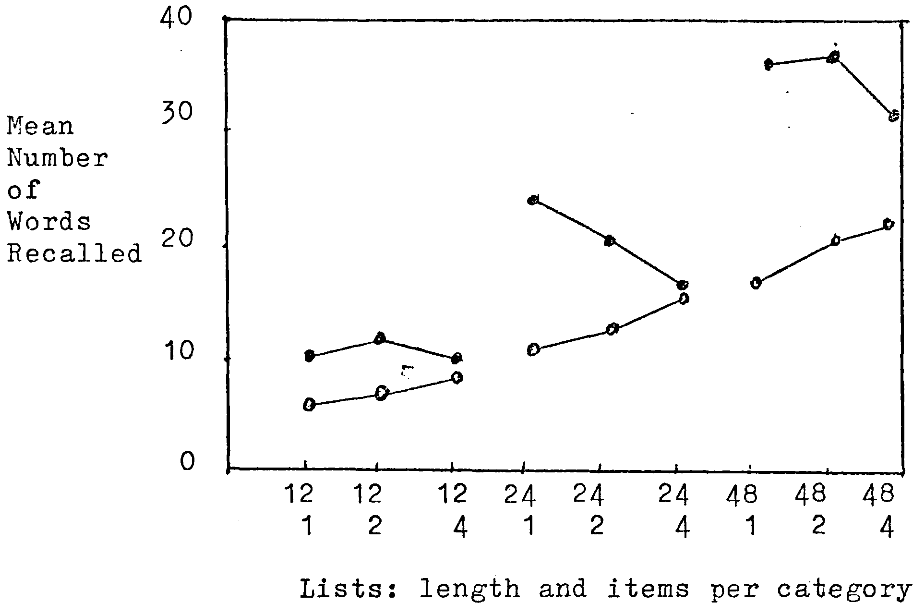
Kintsch used a method of analysis of clustering developed by Johnson (1967) which identifies clusters in recall in the order of their intercluster similarity. The clustering is defined first in terms of the adjacency of two items (symmetrically) and then in terms of the group of words with the highest values of adjacency. The clusters are then extracted in hierarchical form with the most closely related being identified first. The model predicts if the similarity relations among a whole set of output items are evaluated via the adjacency method, they should be related to the underlying semantic structure. Results of an experiment reported tend to support this in that the clusters extracted agreed with the objective semantic structure of the items presented. Kintsch admits that this is only qualitative support for his model but it can be seen as a useful attempt to take the investigation a step further. One important reservation that he makes concerning this type of analysis is that by its very nature it will only illustrate the structure which is common to all subjects - actually what determines the clusters is each individual's idiosyncratic organization. Individual differences appear as noise in the analysis (the degree of noise can be estimated by comparing two methods of analysis suggested by Johnson).

The contribution of this type of model is obvious. It can be seen as an elaboration of the approach embodied by Morton, and indeed is to a large extent compatible with it. On the basis of such new concepts as search and organization it makes predictions which seem at least qualitatively supported by data such as Miller (1967), Cohen (1966) and Tulving and Pearlstone (1966).

Shiffrin (1970) also deals with the search process in a development of Atkinson and Shiffrin (1967) and Shiffrin and Atkinson (1969). Here the search process is a recursive sampling of Image units (I-units) in the LTS, the units having the same type of existence as the stimulus elements in Estes' Stimulus Sampling Theory (Estes, 1959). The recovery of these I-units allows a series of decisions, like whether to continue search, or whether to output the item. As each sample of I-units involves I-units from overlapping sets, the system is not unlike that proposed by Kintsch. As with Kintsch, the items appearing in an experiment are stored with a temporal cue, which aids search and in fact the search set will be delimited by these cues. The nature of these temporal cues is unspecified and this could be a difficulty since it leaves a tape recorder analogy where the search is conducted by looking at the appropriate piece of time based information.

The data cited does support the predictions of the model in the field of free recall and it is clear that there are grounds for postulating a search process such as this. In application to studies of categorization (such as Cohen, 1963 and Tulving and Pearlstone, 1966) retrieval has two fairly independent processes - one concerned with the recall of categories and one with recall of words in the category. The probability of choosing information relating to a particular word or category on a draw is the ratio of the strength of the word or category to the strengths of the entire list available. It quite clearly agrees with findings such as Tulving and Pearlstone (1966) where cuing the category name increases recall and the advantage of cuing increases with list length. The model also predicts the unusual finding that within a certain list length, for the non-cued recall performance rises as category size increases, whereas for cued recall performance decreases as category size increases. In non-cued recall this is explained by the fact that as category size increases a greater proportion of retrieval time is spent in the relatively efficient category searches than in the search for individual words, while in cued recall the increased category size makes category search less efficient. Other support is obtained from the fact that words from the same category tend to be output together. The actual fit of the model is quite close as can be seen from figure 2 (Shiffrin in Norman page 419).

Thus it would seem that there is some future for models of the search process in the study of retrieval from memory, whether one prefers Kintsch or Shiffrin. Problems exist in that models discussed in this section are all involved in the structure of recall - on average i.e. retrieval of relevant words from groups of subjects rather than retrieval of specific words in individuals. Though a strength hypothesis can be made to determine the order of recall or the probability of retrieving a specific word as a ratio of the total strengths,



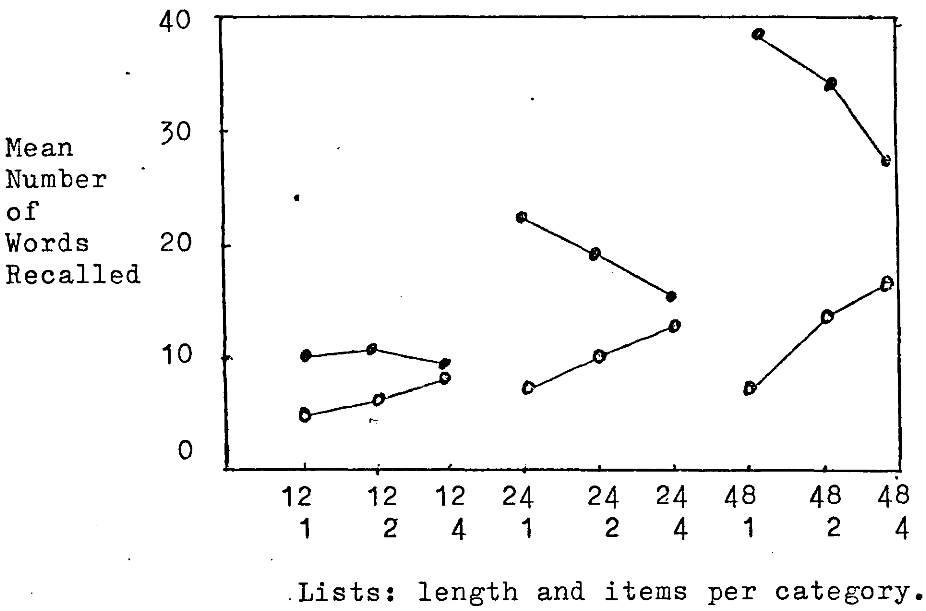


Figure 2.

Mean Number of words recalled in the first recall test as a function of list length (upper number on abcissa) and number of words per category (lower number on abcissa) for testing either with or without category names provided.

(a) upper panel - observed data ● cued recall

(b) lower panel - predicted values o non-cued recall

(Shiffrin, 1970, p. 419)

there would seem to be no way of specifying these strengths for any one individual. Thus in some respects in fitting our model to group data we ignore the essence of the organization which is individual idiosyncrasy. Despite this the idea of incorporating the search process as a fundamental mechanism, is a very sound one in information processing terms and as should later be seen, a very useful one in the study of organization in memory.

RECENT RESEARCH

(iii) In this section more recent developments of the approaches mentioned above will be considered. Having tacitly accepted the type of system proposed by such as Morton (1968), research interests moved on to further thoughts on the organizational processes and how they aid retrieval.

One interesting strand of research was the distinction first made by Tulving and Pearlstone (1966) when they tackled the question of availability versus accessibility in long term storage. Their findings were that there was greater recall under conditions of cued recall than under non-cued recall. This suggests "that specific information about many words must be available in the storage, in a form sufficient for the reproduction of words, even when this information is not accessible under a given set of recall conditions."

More recently Tulving and Thompson (1971) have shown that the retrieval cue has to be encoded along with the to-be-recalled event to be effective. However an interesting problem presents itself here since though two words are encoded at the same time and one can serve as a retrieval cue for the other, it does not necessarily follow that the two words are stored together (as one would expect from an associative hypothesis). Slamecka (1968, 1969) reasoned that if words are stored together then list words should serve as an effective retrieval cue for other list words. Since in his studies, cued subjects did not recall more words than non-cued subjects, Slamecka concluded that the storage of items is probably independent. This was supported by Freund and Underwood (1969) but not by Allen (1969) and Hudson and Austin (1970), who showed that in certain circumstances facilitative effects accrue. A basic mistake here might be that the question is not one of all-or-none recall of words but rather partial cues producing the word with a certain probability and thus leading to such effects as "tip-of-the-tongue" phenomenon.

However there are some methodological arguments against the acceptance of Slamecka's findings. Since the lists he used were small (only 30 words), the number of memory units recalled may have been the same as or greater than, the number formed. Therefore if all the units were available there is little reason to expect the list cues to increase recall.

The upshot of this research is that no firm statement can be made as to the structure of storage though a principle of simplicity might suggest the associative or dependent type of structure as most likely to conform to published findings.

A related problem also has confusing findings. This concerns whether the cues or higher order memory units constitute a different, higher level of storage from the ordinary unit (implied by Collins and Quillian, 1969). Wood (1971) investigated whether the process of forming large memory units was the same as forming small memory units, by studying the effects of forcing subjects to reorganise large memory units and also small memory units - the same negative effect was found, suggesting that the process of forming higher order units does not constitute a different level of storage. However based on the subjective reports of their subjects, Matthews and Manasse (1970) illustrated differential use of category labels in recall of word clusters. Other evidence by Dong and Kintsch (1968) supports this view that subjects can give labels for the memory units they form and the presence of these labels at recall facilitates recall and that the label may be stored along with the appropriate cluster, though dropped at recall. At the same time Hudson and Austin (1970) suggest that there is no evidence that category labels are more effective retrieval cues for stored memory units than elements of the memory units.

Again these opposing views may both be true if attributes of the events are stored as part of the encoding process. The same type of attributes will be formed no matter what the size of the unit to be formed and the attributes will function as partial retrieval cues for the items in the units defined by the attributes specified. Norman and Rumelhart (1969) specify extraction of the attributes as part of the recognition process, and Underwood (1969) offers evidence to suggest that many different attributes of the to-be-recalled material may be stored. The importance of a particular attribute is believed to be dependent on the nature of the task and the material e.g. as meaningfulness increases associative verbal attributes become more important and acoustic less important. The problem of assessing the relative importance of various attributes becomes very difficult as they may not be independent e.g. if the subject remembers the acoustic attributes he may be able to reconstruct the orthographic attributes.

In fact through all this confusion we may be coming to a more balanced view of the working of memory, in that the system that could explain the above findings suggests a system of attributes supplying partial cues in recall, the nature of the attributes determined by the situation and the task requirements.

A further implication of the storage-retrieval distinction has manifested itself in the study of recognition memory. The problem hinges on the hypothesis that recognition involves only storage, or perhaps a minimal amount of retrieval while the process of recall involves both retrieval and storage. This view is fairly generally held (Kintsch, 1970a and Wickelgren, 1970) with the recognition process being described in signal detection terms (Lockhart and Murdock, 1970). McCormack (1972) provides a useful review of the findings in this field, and essentially presents the evidence which supports this view.

A good deal of the research quoted revolves round the assumption that retrieval embodies the organizing part of memory and that storage does not, and that therefore organization of the to-be-remembered material should help recall but leave recognition unaffected. Research by Dornbush and Winnick (1967) and Estes and Da Polito (1967) supports this. The difficulty is however, as before; what kind of storage is there - is it organised in some fashion? If there is organization in storage the recognition-recall distinction does not hold up in the way presented, since both processes involve entry into the storage system and would therefore reflect the effects of organisation. This question becomes similar to the one of dependency in storage and here it could be held by an associative theory that there is organization in the lexicon itself. One way out of this dilemma is to change the nature of the problem by assuming an attribute extraction process as the basis of encoding (Bower, 1967) - thus the problem of recognition becomes one of identifying a specific attribute (in most cases some type of time marker or familiarity index) given to an item whereas recall becomes a problem of identifying an item given the attribute or dependency. As mentioned above Underwood (1969) supports this type of approach.

Strangely McCormack (1972) discusses the word frequency paradox and the research generated by it as supportive of his position. The whole problem is treated in much greater depth by Gregg (1974) though again the solution in terms of retrieval is not forthcoming. The point about this research is that it can be used to give credence to the approach suggested above - that recognition is the process of identifying an attribute. The word frequency paradox says that high frequency words are easier to recall than low frequency words but that low frequency words are easier to recognise than high frequency words, when subjects are asked to decide whether a specific word had already been presented. For recall, since there is a greater number of attributes with high frequency words these words should be more easily specified and therefore more easily remembered, than low frequency; with recognition, the problem is a slightly different one - the attribute required is familiarity, but in this instance it is made up of two conflicting influences: frequency of usage and recency. With high frequency words these two attributes are confused in recognition - the identification of the attribute is hampered by the lack of a suitable code which unambiguously indicates recency - whilst in low frequency words this is not a problem to the same extent, as frequency of usage does not contribute greatly to the familiarity judgement.

This use of codes based on feature extraction is increasingly becoming the popular view in present-day thinking.

One further word on the study of recognition - the work summarised by Mandler (1972) indicates that the original simple distinction of recognition as involving no retrieval, is not tenable. He discusses his experiments (Mandler et al., 1969) where the findings have been that organisation does exist in recognition. The experiment is a test of the hypotheses set out above and is intended to illustrate the point that recognition on the basis of familiarity is no different from that on the basis of class recognition, (very similar to what has been called identification, above). Mandler basically examined different types of confusion arising in recognition and showed that these varied with differing levels of organisation of the to-be-remembered material. Discriminability of an item decreased with decrease in the organisation allowed. He also found an increase in the false alarm rate with decrease in the organisation which would be predicted by a search for attributes, incompletely specified. Also discrimination between the two lists used in the experiment was not affected by the degree of organisation, which could be interpreted as support for the minimal retrieval hypothesis but is probably better used to suggest that though there exists a process simply to measure what Mandler calls occurrence tags this is not the whole of the recognition process. It would also allow an explanation of the findings of Kintsch (1968), in that specific occurrence tags or attributes of familiarity, associated with particular lists and by inference, particular conditions may be unaffected by organisation.

Two further strands of research of interest are embodied in the work of Millward and of Schulman.

Corbett, Rice and Millward (1974) investigated the reaction time of a subject's verification time that a word is a member of a category denoted by another word, as a function of (1) category size (2) organisation of the category (3) production rank of the word as a response to the category name (this measure was taken from a task of generation prior to the measurement of RT). Research previous to this (Landauer and Freedman, 1968; Meyer, 1970, and Wilkins, 1971) suggests that RT in identifying a word as a category member is a function of the size of the category. However there are some difficulties in these studies in that the definition of category size was often based on the principle of set inclusion (since all dogs are animals it is assumed that animals are a larger set than dogs). Collins and Quillian (1970) point out that the differences in RT may simply be artifacts of the nested sets. The experiment described by Corbett et al. allows the subject to define his own category set and the words to be verified are taken from fixed positions in his generation of the exemplars of the category. This is a particularly useful refinement since it allows a group measure of the category size from the total number of responses produced by the subjects, and at the same time utilises the words in the subject's personal system of organisation as the test.

Their findings were first that RTs for large categories were greater than for small categories, both for positive and negative items; second, for positive items large I (a sign of category organisation) produced faster RTs, but no difference for negative items; also low rank or dominance produced faster RTs.

Their explanation of this is in terms of the organisation of semantic space:- Associated with each category and item name is a list of retrieval cues. These cues serve to direct the search usually to what is called the core meaning area, which in turn elicits the best examples of the core. In a production task the examples are produced first and then the next example is generated as a combination of the features of the first and the second item. In this way the Subject generates contingent examples. This is visualised as being a type of random walk through the possibilities (Millward, 1973).

One limitation here is the fact that the approach allows consideration only of semantic space organisation and so does not completely satisfy the coding approach as seen by previous Psychologists (Morton, 1969, Conrad, 1970) and as will be detailed in the next section. That is, it seems probable that there are other types of attributes which differentiate among the items.

Corbett et al's explanation of the verification results involves the assumption that there are 2 sub-processes involved: retrieval of the information relating to the category and to the item and comparison of the two. Category size and organisation affect the retrieval and the comparison is affected by the dominance. These assumptions can be questioned but their acceptance invites the following explanation.

With large categories there are simply more words to be searched and therefore retrieval takes longer - with negative items the search has to go on relatively longer even if there is a stopping rule. The RT is correspondingly longer for larger category verification.

With more organisation the search can be more systematic and therefore faster. This however has certain problems, basically because of the definition of organisation in this instance. It is identified as the average number of items a subject produces when given a particular category name. Is it not more likely that this gives an index of how accessible a set of words is - and then almost certainly how they can be retrieved in verification i.e. the measure seems to involve some circularity. No effect of negative instances would be found as the fluctuations in RT would be random in relation to this variable.

The comparison process relies on the degree of overlap of the two sets of attributes, for the category and the item, and this overlap will be greater with items which are of high dominance. Therefore there will be a faster RT. Again here there may be a confusion with a variable of accessibility in word space. A simple distance assumption may solve the problem.

This paper confirms the trend towards seeking the solution to memory problems in terms of the attributes of the items stored. It is rather more important here as it examines the subject's ability to generate items given some key, which is after all what may be involved in an open-ended situation allowing creativity.

Schulman too, (1971 and 1973) is indicative of this approach embodying the attributes involved in the encoding process in memory. Some of his findings indicate that memory for semantic relations surpasses memory for structural relations. Deciding that a word contains the letter A or that it has a repeated letter, provides a poorer basis for its later recognition than deciding that it belongs to a certain taxonomic category. This is confirmed by Craik and Lockhart (1972). His thesis is that in order to study how experience is encoded, we should construct a range of memory tasks so that the content and the structure of the encoding may be inferred from the results.

MODELS 2

(iv) Further research has led to more concrete conceptualisations of the memory processes. Notable among the theories proposed are those of Posner, of Estes and of Bower.

Posner and Warren (1972) set out 4 propositions:

(1) each stage of encoding gives rise to a memory appropriate to the code created at that stage;
(2) new codes do not obliterate previous old codes;
(3) retention of each code depends on the amount of concentration that a subject invests on a given code;
(4) the memories that represent the various codes may be differentially available to different types of retrieval operations.

On the basis of previous work (Posner 1969) 3 general types of code are visualised - a physical code (involving information on the physical characteristics of the stimulus) a name code (some internal form of identification) and a semantic code (placing an item into meaningful relation with the other contents of memory). The implication here is that this is a natural process of analysis and that these codes are available at different points in the coding process. Also something must remain of these encodings in long term storage. The authors also reject the notion of identifying types of code with temporal parameters, which appears in a great deal of previous research, since certain studies have shown that physical codes are available for much longer than research had initially found (Coltheart and Allard, 1971) and since one can learn and identify things like accents (presumably on the basis of their physical characteristics).

They confirm the previously mentioned dissatisfaction with the trace approach in that some form of differential tagging is necessary in order to encode certain attributes. Hintzman (1970) and Hintzman and Block (1970) indicate that the tagging idea is not a useful explanation since the subject does not know which aspects of the stimulus are going to be tested.

What this system seems to postulate is a structured encoding system, more structured than Underwood (1969)'s suggestions and almost in line with the system Morton (1969) has proposed without the strict time base of encoding. In fact, the name code is closely linked to a Logogen system, acting as a key to further information in memory. The value of the name code is probably only in its similarity to a Logogen, as an internal representation of the unity of the stimulus, since Posner's work (Posner and Boies, 1971, and Posner and Klein, 1971) has been done with single letters as the to-be-remembered items and these are obviously semantically representable as a name. However, if ordinary words are used as the stimulus, the need for and the actual existence of a name code, separate from the physical and semantic codes, may be seriously questioned. On the whole the ideas expressed are a viable system of coding structure and the distinction of physical and semantic codes is used in the work to be reported.

The approach of Estes (1972) is important for a slightly different reason. What his paper sets out to do is to bring together the two streams of thought of association and of encoding. While accepting that what is stored in memory must be some coded version of the stimulus input, he expresses some reservations concerning the use of the idea of coding and decoding as used by such researchers as Johnson (1969), where the decoding process while being incompletely specified, tends to indicate a complex controlling process which in the end may be little different from the confusions of strict associations.

His model suggests the use of a control element which is associated to the items which would in normal association terms be associated with one another. These control elements are in turn related to one another and so on in a hierarchical structure. These control elements may be taken to represent attributes which provide the link between the items joined by the control element. Indeed, Estes specifically compares his system with the feature extraction of Norman and Rumelhart (1970) and the Images dealt with by Shiffrin (1970). The system works as follows: at the time of input of the to-be-remembered items, some element or aspect of the current context serves as a temporary control element. This situation exists while the items are in short term memory, and are being rehearsed - stability is obtained if some item in long term memory is activated by the joint effect of the existing associations with some aspect of the context and one or more of the input items.

Retrieval occurs when the control elements are associated with cues which constitute part of the context at the time of recall. One interesting aspect which is hinted at by the structure of the recall system, is the use of a motivational element as the initial step in the search for the item. This opens a way into problems in Cognitive Psychology and the conscious processes involved in recall.

Much of his work is based on the investigation of order effects in STM and therefore is not important here, but the implications for LTM are interesting and the value derives from his awareness of the inadequacies of both the simple association hypothesis and the unspecified coding approach.

One final and most recent model which attempts to incorporate many of the features mentioned previously, is that of Anderson and Bower (1973). Essentially this is a comprehensive computer model of the linguistic processes, based on the recognition and memory for, grammatical structure. Thus the model perceives and stores aspects of discourse, rather than lists of words, or features of words. The lay-out of the model is shown in figure 3. Briefly information is registered by sensory registers recoded into higher-order features and held in limited capacity auditory or visual buffers. The parsers analyse the contents of these buffers and produce a description suitable for transfer to long term storage.

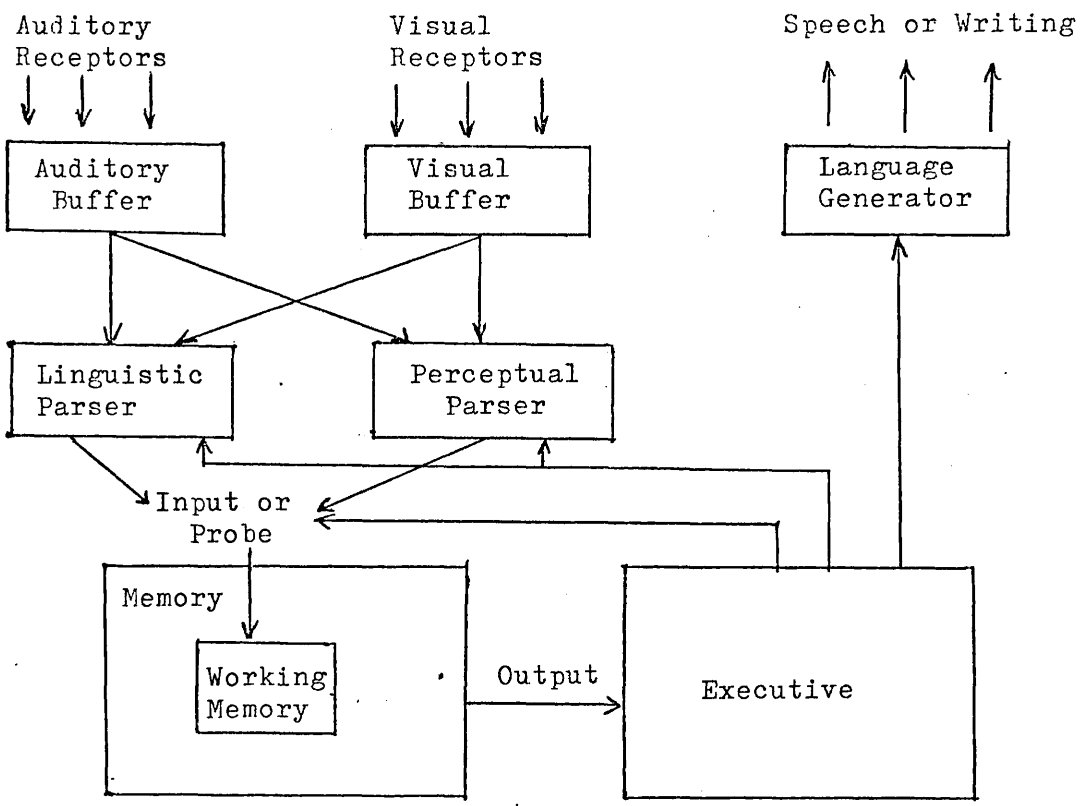


Figure 3. HAM (Anderson and Bower, 1973, p.137)

The linguistic parser translates natural language statements into conceptual descriptions, while the perceptual parser builds up a description of the sensory contents of the buffers. The output of the parser is sent as a probe to be matched to the appropriate contents of long term memory. The probe and its matching structure is output to the executive, which controls all the information processing - leading to further probes, storage or to output.

The system is very ambitious and deals with a great many complex problems in the recognition and storage of natural language. It is unfortunately over-reliant on computer language and processes, but there is much which is of interest here.

The model does no more than acknowledge the physical codes produced in the initial stages of analysis (but they are similar to Posner's ideas and Morton's) and concentrates on the characteristics stored later.

The name of the model, Human Associative Memory (HAM) derives from the associative relations built up in the tree structures generated by natural language. This is clearer when they make the distinction between the idea node and word node, for each word input. The idea nodes are essentially nameless entities that acquire their meaning from the configuration of associations into which they enter with other ideas. The actual English word, is a separate node through which the orthographic and articulatory parameters are accessible. This almost seems to hark back to the learning models, whose meaning exists only in the pattern formed, except in the existence of an executive. The word node is connected by a member relation to the concept of the word. Intuitively this is very convenient, since it would allow tip of the tongue phenomenon where the name cannot be retrieved despite other information being available.

In encoding the parser calls upon 3 types of storage; firstly the auditory and visual buffers, secondly a push-down store of context information, to indicate the context of previous items, and thirdly a working memory, much the same as Atkinson and Shiffrin's (1968) buffer, for transferring information to long term store. Thus HAM does incorporate many of the encoding findings in its input structure and for these reasons does explain many of the memory findings.

Within long term storage the search is a quasi-parallel simultaneous search from all the nodes of the input tree. Each node sends out a sequential matching probe (hence the quasi-parallel) and the probe attempts to obtain contextual information relevant to the item. The executive will then decide which aspect of the retrieved information is relevant for the purpose, and thus will have identified the input. In obtaining fact information say, concerning the use of the word, HAM will confine itself to the required relation, neglecting others (for instance, superset information) HAM selects the best matching memory structure to the k-elements of the probe, using a threshold to decide how many element matches are required for output of the correct signal.

The value of HAM cannot as yet be assessed, though it is clear that it will account for a large portion of the findings mentioned here, if we add the clearer opinions on the sensory characteristics, attributed to Posner and Warren (1972). It is similar to Kiss's ideas, though it is much more comprehensive in its approach to the organisation of memory and the input and output controls. Again it also attempts to combine the associative and the coding approaches, as in Estes (1972), and this in fact is the current advance being made in the field of memory studies - an attempt to come to terms with the notion that both association and encoding are essential as basic blocks of memory.

The information processing approach, in general, has progressed very rapidly from the initial concern over temporal parameters and capacities of parts of memory, to a more realistic appraisal of the organisational elements in memory and just how their effects can be explored.

FURTHER APPROACHES

To complete the section on memory it may be useful to examine two approaches which give an indication of the way memory research will develop in the next few years. The first derived from the work of Pribram (1971) arises out of a long series of experiments concerning physiological psychology, the second follows on from the information processing approach and is seen in the work of Meyer (1973).

Going back to the point make by Kiss (1971): though as Psychologists, we should not seek reductionist solutions to the problems encountered, it would be foolish to ignore the developments made in the field of physiology and not to structure our psychological models in the light of these facts. In particular, findings that it is possible to destroy large areas of the brain and still have little or no effect on highly complex behaviour, cast some doubts on the emphasis of information processing models on fixed structural systems. The idea of single searches through memory, such as Millward's random walk, may be contrary to physiological knowledge -

> "Since the convergence of many impulses on any one neuron is required to make it discharge, chains of single neurons cannot propagate a wave of activity through the cortex. Rather the propagation resembles an advancing front of multilane traffice, with many cells activated in parallel at each synaptic linkage in the chain...." Eccles, Scientific American, 1958.

There have been a number of attempts to characterize learning and memory in some physiological fashion (e.g. neural growth, functions of glia cells, and most notably, the RNA and protein studies) with very limited success and certainly with little insight from the psychologist's point of view. Pribram (1971) has suggested a system which does open up interesting possibilities from the psychological stance. With the TOTE mechanism as fundamental he confirms the idea of feature detectors and analyzers, producing interacting patterns of excitation which as they are transmitted, become organized in parallel. In some measure the feature detectors may produce some sensory code. The most interesting idea appearing in this research is that a hologram may form the basis of some long term storage mechanism.

The hologram is a recently discovered optical phenomenon. It is the storing on a filter of the wave patterns produced from an object, by illuminating it with a coherent light source. When the process is reactivated, the waves can be read out to produce an image of the object, i.e. when one looks at the filter illuminated from behind, a 3-dimensional image is seen. An important point is that there is no image inherent in the hologram itself, the image exists in space in front of the hologram. While being an interesting scientific toy it has very interesting implications. It can be described mathematically (therefore the hologram need not be based on a light source as long as the mathematical laws are complied with) and Rodieck (1969) has shown that these mathematical requirements are satisfied in the shape of the visual receptive field. Thus notions of lateral inhibition in the visual receptive field can be described as patterns of interference and the output from it may correspond to some holographic function.

Given that there is some basis for these structures in one part of the neural system, it may be acceptable to look for the structures elsewhere.

The physical hologram has many interesting features which are similar to a memory system - the information is not localised but distributed throughout the plate and the image to be produced is resistant to damage to parts of the plate. Given a tiny part of the original record all its contents can be reproduced. The capacity of the record has huge parameters; the wave length of the light used which is variable over a very large range, each point accessing a different wave pattern, and also the angle of the light source stores and retrieves different information.

What Pribram proposes is that the hologram may exist in the arrival of impulses at neuronal junctions which activate horizontal cell inhibitory interactions. When the arrival patterns converge interference patterns arise and this interference is distributed over the entire extent of the neural pattern. If we assume that this activity can be made more permanent by some neural facilitation, then we have a model of long term storage. Retrieval from this demands only the repetition of the pattern (or the essential parts) which originally initiated storage. (The findings of Tulving and Thompson, 1970, where the cue and the to-be-recalled item must be stored together for the one to act as a cue for the other). This type of system does away with the need for location information since the required information can be directly accessed by providing part of the context.

The idea of using interference patterns is not a new one, Willshaw, Buneman and Longuet-Higgins, 1969 proposed a similar type of system though specifically not based on the hologram. For the psychologist the hologram seems the more attractive system as it does give the required flexibility to memory, allowing the developmental acquisition of context-linked holographic information. All the codes previously discussed would be held in the one image, and only part of a context would be required to produce all the information needed for complex judgements or for straightforward recall. Anderson and Bower's executive would be required to interpret the pattern and to feed back commands for further search etc. At present the system must be regarded as highly speculative but as it represents developments at an interface with psychological memory research, it is certainly of interest.

More psychological in its reference has been the work of Meyer (most recently 1973) being based from the standpoint of a broad interpretation of the information processing approach. At the same time his ideas may be linked in some sense to the above through his invocation of more basic ideas of excitation, in the traditional sense. His work follows the approach of Sternberg (1969) in the use of Reaction Time as an indicator of more fundamental processes in the memory system. He has investigated mainly semantic memory but the ideas can be equally applied to other aspects of long term storage. Characteristically, the experiment reported (Meyer, 1970, Meyer and Schvaneveldt, 1971) has been the presentation of two key words about the relation of which the subject must make a decision. This RT has been measured as a function of the actual relation of the two words, as a function of the decision made by the subject and in terms of the varying category sizes of the two items under study.

Initially the findings concerning a to-be-verified superset relation of the form "all S are P", indicated the necessity of a two stage model of decision: whether S and P are related and then how they are related. It is interesting to see that Anderson and Bower's model would handle this type of data in much the same way, as the Executive would establish by association that the two words were related, and then by a new search that one entry produced all the others but not vice versa.

In postulating a distance model in the organization of semantic memory, Meyer and Schvaneveldt (1971) suggest a physiological mechanism of spreading excitation as an indicator of the degree of facilitation gained by items on the retrieval of items close in semantic distance. The credibility of this is supported by the findings that when two series of letters are presented and the subject must indicate whether they are words the time taken to access the information for the second decision varies directly with the degree of association of the two items; and hence by implication with the distance apart of the items. The model, supported by Meyer, Schvaneveldt and Ruddy (1972) states:

> "... retrieving information from a particular memory location produces a passive 'spread of excitation' to other nearby locations facilitating later retrieval from them."

This type of concept is not the same as Pribram's wave patterns though the notion of facilitation may be in common, but the importance in the concept is that it indicates the need for an interface with notions other than those directly derived from computer investigations. Meyer sticks quite closely to the data base idea of Kiss etc. and memory is seen as nothing more than a series of storage locations, but principles such as this open important aspects of a dynamic memory system.

Meyer, Schvaneveldt and Ruddy (1972) give evidence that increasing the interval between two associated words decreases the facilitation of the RT for their identification as words, whereas separating them with an unassociated word does not eliminate the effect of facilitation in RT. Associated words are judged faster than unassociated words. Spreading excitation is shown to be a better explanation of these findings than location shifting, which is the natural theory of the coding hypothesis (this assumes that given the code the word can be retrieved without disturbing the other words in the store). The important conclusion is that it allows us to characterise facilitation of context, since the retrieval of the first item creates a context of lowered thresholds (by the spreading wave) and this context allows easier retrieval of the second item from that memory area. Strangely enough this is almost the result implied by Pribram's ideas, though the crucial difference is that when an item is retrieved the <u>whole context</u> is brought to the executive part for selection, and consequently the items closely related to the chosen (or verified item) will more easily be recognised in further tests.

The intention in quoting these two approaches to the problem of human memory is to illustrate how diverse fields will contribute to the study, and why it may be necessary to invoke principles not directly attributable to information processing, which is increasingly relying on computer terminology and expertise as a proving ground for the theory.

The conclusions to be drawn from this review of the field of memory research are two fold:

first, the early approaches which stressed the capacity and the temporal parameters have to a large extent been superceded by a concern for the nature of organization in memory, mainly in terms of how input is encoded and then how representations are retrieved;

second, that previous distinctions between association principles and the doctrine of coding cannot now be usefully upheld as separate theories of memory. It seems inevitable that any theory of memory must involve the use of codes which mirror the way we perceive our world, yet at the same time there is a need in memory for some mechanism of dependence whereby items are contiguously arranged by importance - this is probably best described by the principle of association.

It is not clear where the mainstream of this research will lead but it may well be the case that individual differences and correlations with other cognitive abilities dependent on memory, will shed an amount of light on the study. This line of argument will be examined in this thesis.

Research involving Creativity and Memory

This final section should serve to indicate the ground work on which the approach adopted in this report, is based and from where it derives its motive. However in this instance there is a dearth of research which involves both the fields which are under discussion. One study which set out to investigate creativity through the techniques of learning and memory is that of Maier, published in five separate reports but b asically involving the same methodology. (Maier, Julius and Thurber 1967 I, Maier and Burke, 1968 II, Maier et al., 1968 III, Maier and Thurber, 1968 IV and Maier et al. 1968 V).

They suggest two methods of approach to the problem of creativity: 1. the creative mechanism is the same as for learning and therefore represents a particular aspect of the learning function, 2. creativity is a higher mental process than learning - a gift some people have. They began from the former viewpoint. The type of question they ask is "... whether some people are relatively more inclined to use elements in thinking that have been associated in past experience, whereas others are relatively more inclined to neglect associative bonds or to make new combinations of old elements." I.(p.494)

Their experiments consisted of allowing the subjects to learn paired associates to a certain criterion, and then asking them to construct a story. There are three types of paired associate relationships:

1. neutral - cannot easily be used in proximity (diamonds-warden)
2. congruous - fit together in the story assigned (wicked-drunkard)
3. incongruous - do not fit together in the story assigned (dry-basement)

The stories are then scored for originality as in Getzels and Jackson's (1962) study. Unfortunately there was no clear relation between originality and particular word usage, though their theory would have predicted that fragmentation and recombining would be related to originality. However they did find that fragmentation was a consistent ability and therefore worth further study.

When they encouraged subjects to write creative stories -

> "It is apparent that reducing the preference for old pairs reduces the frequency with which they are used but does not eliminate them. Furthermore, instructions failed to increase the creation of new pairs ... " II. (p.357)

They went on to investigate whether degree of learning influenced the type of usage.

> "One might expect the strengthening of associative bonds to increase the usage of old pairs and decrease the usage of single words and new pairs... The fact that neither occurred, despite the memory score's significant improvement, and the introduction of over-learning, suggests that reinforcement might operate

to increase recall, but that in a problem solving situation some kind of selection process operates to offset the strength of the associative bonds." III (p.367)

Perhaps the type of search adopted by the subject in this open ended task involves the use of codes inappropriate to the previous association. A further experiment tends to confirm this. When organization in recall was studied prior to the story writing -

"Clustering was evident but the stability of clusters varied. However when the problem of using the words in a story was presented, no individual used the words according to the way in which they were stored. Even those with highly stable recall lists completely abandoned the organization established in recall." V. (p.1020)

Despite the general difficulty of the series of studies in that they did not measure what they set out to, they do illustrate the point that meaningful things can be said about cognitive processes with the established techniques of memory research. The actual findings are interesting in that they suggest that subjects use different search strategies according to the task and there is a slight implication that use of paired associates may not indicate the normal usage of memory in verbal tasks.

EXPERIMENTAL APPROACH

The difficulty with the previously mentioned studies of Maier et al. (1968) is that they lost sight of the creativity link and failed to establish the measure studied as an important consistent ability. While it is difficult to determine just what the ability measured was, it is relevant to note that the factors of flexibility outlined by Guilford (1967) and discussed by Wallach (1972) correspond in many respects to that of fragmentation, in their seeming importance in creative thinking but their close affinity to convergent abilities. The general approach of Maier is important however, in that he used an experimental measure to tap a cognitive ability.

The approach adopted here is similar in that it is hoped to characterise a hitherto known cognitive ability as an individual variable worthy of experimental investigation, in the manner of the study of memory. The idea behind such a mixture derives from work such as, that of DeGroot (1965) when he studied the recall of Chess masters of complex middle game situations exposed for a brief period. The difference between the recall of masters and that of competent chess players, apart from a quantitative one, where masters recalled almost all of the situation while ordinary players recalled much less, was in the quality of the recall. The masters tended to impose an organization of their own on the material which reflected their experience and feelings, essentially a dynamic coding, while the other players tended to attempt a literal organization of the positions with little elaboration in their encoding.

Further suggestions of this approach to the problem of creativity can be seen in Taylor and Barron (1963). In their description of the attributes of the creative scientist, they include the following:

> "The scientist who can respond creatively to a crisis must therefore be of a high order of intellectual ability and must be orderly, thorough and disciplined in his acquisition of current knowledge.
> As discoveries occur which cannot be assimilated to current conceptions of orderliness in nature, increasing effort must be made to understand the unordered and to find a new principle which will restore order. The person who pays close attention to what appears discordant and contradictory and who is challenged by such irregularities is therefore likely to be in the ranks of the revolutionaries."

The clear suggestion is that the difference may lie in the encoding of encountered material and its placement in the structure of previous information.

This is the basic position adopted and what the study set out to do was to indicate any differences in organization in memory, or more accurately the very long term memory base dealing with items the subject has known for a long period, which can be related to differences in creative ability.

The actual methodology varies from experiment to experiment and therefore the details are left for the Introduction of each experiment. The trend in the research is towards refining an approach to the problem rather than following a set of predictions based on an a priori model. Generally speaking, the study falls into two sections: firstly, an examination of tasks which have been traditionally used in the study of creativity and a comparison with other tasks on the same level of measurement - the purpose being to explore the relation of these tasks to general memory performance; and secondly a move to more accurate measures of the internal processes using similar methods to Sternberg (1966) and examining latencies for recognition and classification in tasks appropriate to the processes postulated as underlying the memory system.

Experiment 1

The main idea was to investigate differences in extent of recall corresponding to differences in creative abilities. First, it was necessary to conduct a pilot study into the suitability of various types of materials and forms of instructions. The study involved continuous association, measuring both number of associates and originality of associates. Recall of both stimulus item and association was required on the following day.

4 types of material were piloted:

(a) Glaze (1928) nonsense syllables - one item from each of five association frequencies 0%, 25%, 50%, 75%, 100%.

(b) Consonant syllables (from Underwood and Schulz, 1960) - one each from 5 association frequencies - 0%, 25%, 50%, 75%, 100%.

(c) Dissyllables (listed in Underwood and Schulz, 1960 - not necessarily words) one each from association frequencies - 0%, 25%, 50%, 75%, 100%.

(d) Thorndike-Lorge word count (1944) - an arbitrary selection at approximately equal frequency intervals - one each from - 4 in 18 million, 10 in one million, 30 in one million, A words, 500 most common.

The items were presented in a fixed random order using a slide projector. Subjects were drawn randomly from the student population. Each was tested individually. Each item was seen by the subject, spelt out and then pronounced. The reason for this was to ensure that each subject had access to the same encodings of the items. The subject then wrote down his associations, looking back at the presented item after each association (to protect against simple chaining). One day later, the subject had to recall the item and his associations.

The results can be discussed under 3 headings:

1. the effect of different instructions;
2. which type of materials were most successful in eliciting several associations;
3. whether unusual associations were better recalled.

1. Three sets of instructions were given to different subjects. These were (a) simple association instructions - that the subject write down associations as briefly as possible until he had exhausted all his ideas; (b) high frequency and high originality instructions - the above instructions plus the suggestion that the subject should be as uninhibited in his responses as he could and to write down every idea; (c) low frequency instructions - as in (a) but also requesting that the subject ensure that there was a clear link in his mind between the item and the idea.

Results indicated that there was no difference in the number of associations produced by the instructions, nor was there any difference in the unusualness of the associations produced as a result of the different instructions. The only difference found was that (c) led to better recall of all material. The implication here is that instructions to be original or to produce a larger amount of material are not necessary for this type of task, and that simple instructions on the task will suffice. If recall is desired some instruction concerning the linking of associations may be necessary to increase the amount of recalled material.

2. In terms of the number of associations produced none of the different sets of items followed the association values given in the norms. It is not clear why this should be, other than that the norms are American and are to some extent out of date, though one could argue that language changes relatively slowly. It does confirm the difficulty of using material which has verbal connotations. What is clear is that the dissyllables and the words from the Thorndike-Lorge word count were better associated to and better recalled.

3. Here there were no clear-cut findings, since there was difficulty in rating the originality of the responses; because of the relatively small number of responses gathered. The scoring system used was as follows:

0 - any association which had more than one occurrence

1 - any unique association but whose link with the stimulus can be clearly seen;

2 - any unique association where there is no apparent link with the stimulus, i.e. they are idiosyncratic to the subject.

Analysis indicates that there was a greater percentage of idiosyncratic responses in the subject's recall than in his initial response production. This would suggest that there was some selective recall of unusual material implying tentative support for the broad general hypothesis.

Even at this early stage it is clear that there is an inadequate conceptualisation of the processes underlying the production of associations. The naive approach taken was that when an association was called for, a response was brought from memory and related to the stimulus to form some reinforced bond. This would suggest that stimuli which produced most responses would be better recalled later - however this did not seem to be the case, except with nonsense syllables as stimuli. The answer may be that an association consists of a stored bond which is brought forth as a unit of two words, or more normally, tagged for occurrence. The association does not alter the probability of recall of the stimulus or the response but only the recall of that particular bond. With nonsense syllables presumably there are no existing association bonds and so the words retrieved from the memory store must be paired with the stimulus at that time, and so on for each response. The stimulus comes to be the most rehearsed part of these new associations and is therefore recalled better, than the associates.

In an attempt to throw some light on this problem, it was decided to deal with associations that the subject had to construct at the time of the experiment and would therefore constitute a new bond. At the same time the subject's ability to produce unusual associates to words in a separate task would allow a study of the relation between originality and recall.

The main experiment as a result took the following form:

(a) a free association task, involving continuous association to words - this was to be scored for the fluency of the responses and for their originality on a statistical basis.

(b) subject's production of 10 unusual and 10 common words.

(c) a paired associate task consisting of these produced words being paired with nonsense syllables (since they seemed to be divorced from the normal word memory in the last experiment) of the same association value. The subject generated words which he thought were unfamiliar or familiar, hopefully by-passing the problems of the unsatisfactory nature of the frequency counts of words currently available.

(d) a free recall task of the items in (b), after a delay of 3 hours.

(e) two creativity measures, from the Minnesota tests (Torrance, 1963) included to assess the degree of relationship between the unusual associations and creativity.

AIMS

The aims of this experiment may be expressed as follows:

- to investigate the possible link between the number of unusual associations elicited from subjects and their ability to store and retrieve unusual words when paired with nonsense syllables.
- to study to what extent these measures correlate with conventional measures of creativity.

- generally speaking to indicate the nature of the processes behind these measures in a form conducive to treatment by experimental methods, or to produce suggestions of meaningful refinements for the approach.

SUBJECTS

60 pupils, mean age 15.06 years, from a comprehensive school - taken from the B, C, D and E streams. Subjects were tested in class groups of 33 and 27.

MATERIALS AND DESIGN

The first part of the experiment consisted of presentation to the subjects of 30 words selected randomly from the 2-syllable nouns in the Thorndike-Lorge L count spanning the frequencies 100-333. Certain especially American words were discarded e.g. sidewalk, highway. The words were printed on large white cards in letters 2 inches high and held up to groups of subjects, seated at varying distances up to 18 feet. For each word the subject was instructed to write down continuous associations until told to stop, which was after 40 seconds (this period having been found to correspond to a falling off in response rate, in the pilot study). The complete text of the instructions for all the tests in the experiment can be found in Appendix 1.

The second task involved the subject in writing down on blank cards (4 inches x 2 inches) 10 unusual words, one to each card. These words were defined as words which the subject thought were uncommon, unusual words which they did not see very often. All the words had two syllables and were nouns - this was to prevent length of the word being used as the criterion of unusualness. It made the task more difficult but provided a clearer measure of the unusualness dimension in subjects. The letter 'A' was added in the left hand corner of each card, designating these unusual words as A-words. Following this subjects had to produce 10 common words, everyday words which they came across frequently. Subjects then designated these B-words. The cards were randomized by shuffling. Clearly the A, B system was necessary in order to know which was which after randomization.

Each subject received a booklet of 20 pages, each page having a nonsense syllable printed on the left hand side. Each booklet had the same nonsense syllables but printed in a different random order. These nonsense syllables were taken from Underwood and Schulz (1960), association value around 30. These were then paired with the generated words in the booklet, by writing the word and below it the nonsense syllable, on the right hand side of the page. Instructions were to attempt to see some link between the two items; no mention was made of the fact that they were to be recalled later (in this way it was hoped to come closer to the normal situation whereby items are learned and stored as incidentals and are then of subsequent use).

In the final part there was a creative test booklet comprising 2 pages of circles as in Torrance's Circles test - where instructions are given for the subject to make the circles into a more complex drawing of something which has a circle as the main part - and 2 pages of the consequences test, consisting of 3 questions:

"What would happen if man could be invisible at will?"
"What would happen if a hole could be bored through the earth?"
"What would happen if man could understand the language of the birds and animals?"

Under each, space was provided to allow the subject to make as many responses as possible.

There was a free recall test approximately 3 hours later, where all the words generated plus the nonsense syllables associated to them were to be recalled.

Each group of subjects received the tests in the same order.

PROCEDURE

In the first task, word association the subjects were given the following instructions:

"You are going to see a number of words written on cards. I will hold up each card separately and you will have some time to write down words or phrases which the word on the card makes you think of. After you have written each idea down you should look back up at the word on the card and see whether it reminds you of something else. I will tell you when to stop." There was a 5-minute break after the first 15 items.

In the second task, after issuing the subject with small cards for writing the words on, the instructions were: "On each card I want you to write a word - but there are certain things about these words that you write. First, they must be unusual words, that is words you don't come across very often, and secondly, they must have 2 syllables or 2 parts (demonstrated if it was clear that certain subjects did not understand)."

At the end of the time allotted each person was asked to write the code letter in the left hand corner of the card. The instructions were repeated for common words.

To pair the words produced with the nonsense syllables in the booklet, subjects were instructed after having randomized the order of the words to "place the cards face down on the desk. Turn the first card over and write the word opposite the three letter "word" on page 1 of the booklet. Copy the 3-letter word underneath the word you have just written and try to see some link between the two. When you have linked them in your mind, turn to the next page and turn the next card over. Go on till you have finished the 20 words."

In the final section approximately 3 hours later, (as far as the school timetable would allow) the subjects were requested to recall as many of the paired associate units as possible. This was an attempt to examine whether the associations existed as a unit or were more easily stored as single entities. After 5 minutes of recall they were told that they could write down any word or nonsense syllable, on its own if they could not remember the whole unit. 15 minutes were allowed for recall.

The creativity measures were then administered, with ten minutes being allowed for the circles and ten minutes for the consequences. In each case the subject was asked to think up as many things as possible - no mention was made that they should be original or clever. The instructions were designed according to the rationale of Wallach and Kogan (1965) in order to foster a game-like approach on the part of the subjects. Thus it was presented as an interesting task where they could put down their own ideas.

RESULTS (1)

The correlations are shown in table 1.1. Raw scores appear in appendix 2. The treatment of the data was as follows:
(1) the initial free associations were listed and scored for fluency (the number of associations given by each subject) and unusualness or originality. The statistical frequency of occurrence of each response by each individual was calculated as a function of the group's responses and this value was then assigned as an originality score for each item

Correlation Matrix

Table 1.1.

	Circles	Conseq.	%R-A	%R-B	%R-A+L	%R-B+L	%R-NS
Unusualness Ratio of associations	.23	.26	-.06	-.02	.07	-.04	.08
Circles		.15	-.08	.06	.16	.02	.02
Consequences			.11	-.02	-.20	-.11	-.10
% Recall of A-words				.31*	.36**	-.03	.28*
% Recall of B-words					.38**	.43**	.21
% Recall of A-words + nonsense syll.						.20	.41**
% Recall of B-words + nonsense syll.							.21
% Recall of Nonsense syllable							

* significant at the 5% level

** significant at the 1% level

*** significant at the .1% level

for the subject. The subject's originality score was obtained by summing over all the responses the subject made. The scale used was one suggested by Cropley, 1969, based on equal divisions in the normal curve of .44 standard deviations. Thus responses given by 15 per cent or more of the subjects were given the score 0; responses given by 7-14 per cent, score 1; responses given by 3-6 per cent, score 2; responses given by 1-2 per cent, score 3. The originality scale is a comparison of the subject's responses with his peers, but is a cumbersome scale to operate in its present form, i.e. counting all the different responses made by all the subjects, assigning a scale value to each one and then returning to individual response sheets and computing a score based on these values. However, the advantage of using norms from the subject's present cultural group outweighs the ease of scoring using the test's American norms. A ratio of performance (unusualness x 100/fluency x 3) was calculated for each subject. This gives a measure of the originality "efficiency" i.e. originality per response. This measure can be conveniently called Unusualness ratio for the present. This later becomes Originality Index as a better theoretical justification can be made for it.

Correlations were calculated between the scores for the first 15 stimuli words and the second 15, for the measures, unusualness, and unusualness ratio. Both were high, .6591 and .6678 and significant. The Brown-Spearman coefficient was .7903, showing that the method of scoring for originality is internally consistent. With such an open-ended task as word association the problem is in developing a system of interpretation which consistently reflects the subject's performance. Thus the primary concern here is with the scoring system and by taking a split half correlation and by using the Brown-Spearman formula a measure of the internal consistency of the test may be obtained. Since in this type of test one can assume a certain amount of consistency of responding on the part of the subject, the correlation indicates the usefulness of the scoring system.

(2) recall scores in terms of number correct were obtained for the usual and unusual words and the nonsense syllables, with which they were paired. Five scores were available for each subject - these were recall of unusual word plus nonsense syllable, unusual word on its own, usual word plus nonsense syllable, usual word on its own, and nonsense syllable on its own (since recall of these was too little to separate into unusual and common association).

Certain of the words generated by the subjects were ruled unacceptable for either of two reasons:

(a) because they did not fall into the category stipulated i.e. 2-syllable nouns.

(b) because they were responses involving the same concept e.g. father-mother-brother - only one of these would be accepted. The list of accepted words appears in Appendix 3.

The measure taken for recall was percentage recall, of the acceptable responses. This was justified by the fact that the same criterion of acceptability was applied to the recalled words as to the generated words.

(3) the creativity measures were scored in the same way as the free association responses although in this case only the performance measure or originality index (originality x 100/fluency x 3) was used. This is in line with the findings of Kyle (1970) as mentioned in the review.

All these measures were correlated with each other and the results are shown in table 1.1. The points to note are (i) there is no relationship between unusualness of association and recall of generated words; (ii) there is no clear relationship between creativity measures and the unusualness of associations, despite the fact that both would be expected to have some element of originality; (iii) there is no relationship between creativity and any of the recall scores, so the initial naive hypothesis as mentioned in the first aim, cannot be supported by this data; (iv) surprisingly there is only a slightly significant correlation between the individual measures of recall, i.e. .31 between recall of A-words and recall of B-words, .28 between A-words and nonsense syllables, and .21 between B-words and nonsense syllables. This might indicate different treatments of the different types of stimuli, by the memory system, or that they might be dealt with differently by individual subjects - i.e. common words are not consistently recalled better than uncommon words.

DISCUSSION

The variables of association and creativity seem to have no predictive value for recall. Thus there is no support for the first two hypotheses to be tested (see aims section). There are several other points to be made. Firstly the use of the paired associates to be recalled was not a success, to the extent that subjects had great difficulty recalling the word and its corresponding nonsense syllable together. The rate of correct recall was thus too low to be reliable. Hence, the most realistic measures of recall that can be taken are the individual ones of recall of A-words, of B-words and of nonsense syllables.

Secondly, the correlation between the two measures of creativity is low and non-significant. This finding goes against previous results (Kyle, 1970) where the corresponding relation was .44 and significant at the .01 level. There is no immediate explanation of this though they may be tapping slightly different areas of creative ability, one thinking in shapes and one the use of ideas. Essentially they were chosen from the range of measures as different approaches to the same ability. If it can be shown that they are related in some way, through originality, this will be vindicated.

Overall it is clear that the relation of these ability measures to memory measures is a great deal more complex than was hypothesised. People who generate unusual ideas in one instance do not have a corresponding blanket ability to recall all unusual things. One aspect however, which should be followed up is that stated in point (iv), above, that the lack of correlation between the different measures of word recall may indicate the presence of some uncontrolled variable in memory. The clearest candidate for this is the unusualness of the words generated. It had been hoped that by getting the subject to generate his own unusual words, the degree of unusualness of the stimuli for each subject would have been controlled. However, if there is a variable determining access to the uncommon words which is manifested differently in different individuals, the relative unusualness of the output words for different subjects may be unequal. For this reason, an examination of the unusualness of the generated words was undertaken.

RESULTS (2)

All words generated whether unusual (A) or usual (B) were listed together and originality scores were assigned according to the procedure described in the case of the free associates. Correlations with the other measures were calculated and the results are presented in Table 1.2.

These results were more interesting. This evaluation of the unusualness of the words generated had a significant relation with the other measures of originality and may in fact, represent the underlying common factor. At the same time, the unusualness of the common words was not related to the measures of originality, predicting the low correlation between usual and unusual words produced. This would indicate

Correlation Matrix with the Addition of Unusualness of generated words.

Table 1.2.

	Unusualness Ratio of A-words	Unusualness Ratio of B-words
Unusualness Ratio of associations	.37**	.09
Circles	.36**	.09
Consequences	.53***	.11
% Recall of A-words	-.01	-.04
% Recall of B-words	.08	-.32*
% Recall of A-words + nonsense syll.	-.09	-.15
% Recall of B-words + nonsense syll.	-.03	-.18
% Recall of Nonsense Syllable	-.15	-.12
Unusualness Ratio of A-words		.05

* significant at the 5% level

** significant at the 1% level

*** significant at the .1% level

that those high in originality are not simply producing unusual words when asked for words, which would imply a corresponding high score on common words, but are able to access words at a different criterion level of unusualness. This is supported by a Wald-Wolfowitz Runs test, carried out on the two groups of those highest on originality of unusual words (N = 13) and those lowest on originality of unusual words (N = 13), compared on their values on originality of common words. The use of the Wald-Wolfowitz test here is justified on the grounds that the two groups selected on the basis of a different measure are effectively independent samples. Siegel (1956) says:

> "The Wald-Wolfowitz runs test is applicable when we wish to test the null hypothesis that two independent samples have been drawn from the same population against the alternative hypothesis that the two groups differ in any respect whatsoever." p.136

Thus the test examines central tendency, variability, skewness, any variable which could cause differences. No significant difference was found - indicating that the high group <u>have access</u> to the common words of the others but also to the uncommon words.

Notable of course, is the fact that these originality measures were not correlated with recall scores, thus confirming the lack of support for the initial hypothesis. One interesting figure is that of a significant negative correlation between originality of common words and recall of common words. This supports the common finding that common words are easier to recall than less common words (Gregg 1974); an effect similar to the word frequency effect.

Further tests were carried out, isolating groups high and low on the measures of originality of words generated (N = 13, in each case) but no differences on any of the measures of recall, were found. One final noteworthy feature of Table 2 is the significant correlation between originality of unusual words and the unusualness of free associations. This contributes more evidence on the construct validity of the measure Originality of Unusual words.

DISCUSSION (2)

From these results it can be seen that originality of generated words which are to be unusual, provides a measure which seems to agree with the tests of creativity and also the unusualness of free association. It is not certain from this that it can be concluded that this provides the definitive measure of originality, but it can be accepted that it corresponds markedly to the factor common to all the tests. To this extent it provides a more useful measure of the originality of the subject than any of the other measures on its own.

From the theoretical point of view certain derivations are possible. If it can be conceived that each subject's vocabulary and strategies for retrieval etc. constitutes a "word space", it is quite clear that the original person has not just a different word space, based on a broad criterion difference for production, but a word space which includes the words and strategies available to the less original plus enhanced access to unusual responses as necessary. Thus the original person has a larger word space. This could be due to three variables. Either, he has different strategies available or a larger vocabulary or uses different criterion levels for access. The last two suggestions have been directly studied in a later part of this research, and some inferences can be made about the first mentioned.

It is not possible at this stage to add much on the question of the nature of storage, whether in terms of association or coding. The attempt here to induce storage of associations which were new, was not successful since there was little recall. There could be a number of reasons for this - the task may not have been explicit enough for the subject, the use of nonsense syllables is not so easy for non-University students etc. Probably the best way to conceptualize the memory system is to assume the existence of associational clusters (demonstrated by Pollio, 1964) but allowing a system of coding strategies to cut across the entries in the association net (as visualised by Kiss, 1973) and this would handle the above results.

Originality can be identified with any of the three variables mentioned above, where vocabulary size and criterion may be more closely related to the association type of explanation and strategies to the coding hypothesis. How to separate these experimentally is not yet clear, but it may prove fruitful to control these variables in turn and study the effects on the above measures. This is in fact, the line which this research takes.

CONCLUSIONS

The message from this experiment can only be that the study of cognitive abilities in terms of experimental methods used in memory research, is a complex one. What has usefully been obtained from this initial venture is a measure which relates to the three divergent production measures. While the factor structure cannot be easily disentangled it would seem to provide an efficient indicator of the powers of the other measures. For this reason, it has been decided to study this measure further in an attempt to evaluate its status as a link to the storage system and the processes of retrieval.

Experiment 2

This experiment sets out to examine two of the areas of difference suggested in the last section, namely differences in vocabulary size, leading to the output of more original words simply because the subject has more words altogether and the use of different criteria for the judgement of when an item is unusual. Both these ideas relate to the differences between creative people and others and should throw some light on the reasons for the differences. However to make this meaningful some theoretical framework must be made available for the incorporation of the results. This concerns the question of what type of system is most suitable for characterising the phenomenon of word retrieval when placed in this context. As discussed in the review there are two general possibilities: an association network and a coding model.

The Association model - If the basis of memory is seen as storage of discrete words linked by some laws of association, it is quite clear that the whole is a network through which an individual looks in search of particular word responses. In the normal everyday tasks of language use this will be very simple as words which are used together will tend to be closely linked and therefore their generation will be both speedy and accurate. A search is likely to consist of the excitation of a particular link followed by the testing of the word produced at the other end of the link. This may happen in series or in parallel. Depending on some criterion for usage the word may be generated or used as a further starting point for greater search.

Differences in word usage ability would be seen in the richness of the network. Increasing the number of links between words will make them more accessible and increase the flow of words in language. It is proposed that this type of difference corresponds to a difference in word generation, and that the creative person is identified by an enriched system of associative links. This is intuitively correct, and it also falls in line with the significant correlation between unusualness of word associations and unusual word generation found in the last experiment. It is not clear how a subject decides that a particular item is unusual, though it may be in terms of amount of inhibition encountered in the finding of the item. For the moment the suggestion adopted is that the test part of the system assesses the strength of the item produced and that this threshold for generation may be lowered to allow certain words to appear.

<u>The Coding model</u> - Here, what is assumed is that for each item stored there are a number of codes which specify the word and its location. These may be constructed by the input process or may be inherent in the storage classification. The codes may take the form of information about the meaning of the item, the sound of the item etc., but may also hold information about the familiarity of the item. Generation of unusual words would mean retrieving items according to the code of unfamiliarity, and in many cases this may not constitute enough of the information about the item for the most unfamiliar to be generated. The task of word association,

in the sense that it has been found to relate to word generation, constitutes the extraction of codes from one word (the stimulus) and the use of these in finding other word(s). From this point of view one would expect the 2 tasks to be related. Differences in generation of unusual words then arise because of differences in the ability to use the codes available to retrieve the unusual words. This may be due to experience or training. Again, this is intuitively appealing - for instance differences occur in the ability of certain individuals to complete crossword puzzles (!), mainly due to practice or training, and quite clearly in the poet, whose word retrieval is superior in quantity and subtlety (unpublished work by G. Paul).

In investigating the criteria of unusualness this experiment attempts to distinguish between these two approaches. Following the reasoning that suggests that the difference in recall and recognition tasks arises because the former involves search while the latter reflects the structure of memory only (McCormack, 1972) this experiment presents tasks of a similar nature since the differences in the models centre round the structure. In effect the tasks consisted of first, a word generation of unusual words (retrieval from memory) and second, a rating of unusualness by the subjects, of a selection of words of differing Thorndike-Lorge frequency. The reasoning here is that in the Association model the concept of search, as used by McCormack, is not appropriate and thus to the extent that for this model both recall and recognition involve access to the word in memory, by the same means through the association network, then the ability to generate unusual words should be supported by an ability to accurately rate unusual words. If the notion of the association network as discussed above is examined, the reason for this can be seen. The ability to generate unusual words is determined by the structure or richness of the network. The ability to assess the unusualness of the item presented is also based on the network structure to the degree that the inhibition is only built up through a summation of the links from the word and the strength of the association can only be assessed from the relative accessibility of the word, (in the associative network access must be determined by facilitation and inhibition - the key concepts in the learning approach - thus the extent to which a word is difficult to find is a function of the lack of facilitation or the inhibition of the relevant links). To assume the existence of a frequency counter stored with each word is to assume a code system, and thus an association system requires that the information as to strength or frequency is carried in the associative link. The result of this is that it would be expected that differences in word generation ability will be parallelled by a difference in a rating of unusualness.

The prediction of a coding system is altogether different. Here the ability to retrieve unusual words is dependent on the search mechanism in the form of use of the appropriate code, thus the differences in performance on generation should not be seen in the rating of the words. The task of rating is a different one from generation in the code system - here the word is given and the appropriate code can be extracted by the normal processors of input material. If all individuals have a similar familiarity with language, which we hope for by random sampling from a similar population, they should be equally accurate in rating the words for unusualness.

There is one qualification to this seemingly clear-cut experimental test and it makes the test a weaker one. This arises out of a discussion with Kiss after the experiment. The association network approach may include an assumption which changes the view of the processes at work. This is that the associative links are uni-directional and that what happens in generation is determined by the "indegree" of the item, and in rating what is tested is the "outdegree". These refer to the measure of the linkage available in moving from that item to another (outdegree) or in moving from other items to that particular one (indegree). So what matters for generation is the number of ways in which one can move through the network in order to reach the item, while in rating, the starting point is the item and the measure must be in terms of the number of items reachable from that starting point. The difference in these two processes may be magnified when we have the extreme example of very unusual words.

However despite this being a very important point it need not overshadow the test implied. In the normal circumstances for a great percentage of words the indegree and outdegree will be very similar, and each will be a valid indicator of the use of the word, whether as a "prefix" or a "suffix". With more unusual words this situation could well be different. But if it can be accepted as it has been claimed above, that the richness of the association network determines the ease of access to unusual words for generation, then it is also apparent that it is reflected equally in the indegree and in the outdegree, and consequently, if an individual has difficulty in generating words because of the poverty of his network, by the same token he will be <u>less accurate</u> in his ratings. In conclusion, the test is not invalidated by the assumption of uni-directionality but may be slightly weakened.

The rest of the experiment revolves around the two points mentioned at the beginning of the introduction - namely the estimation of subject's vocabulary size, and whether the creative individual has different criteria, which has been dealt with theoretically above.

The necessity for a vocabulary test is obvious, since a subject cannot be expected to generate words if he does not have sufficient to choose from. It was visualised that there would be some sort of complex relation, similar to the threshold concepts used to apply to the intelligence-creativity distinction, that above a certain level of word knowledge absolute vocabulary size would cease to be an important variable. It was decided after a lengthy search that there was no standard test which set out to give a measure of vocabulary size, as such - all were related to some measure of language use, school achievement or to intelligence. In the course of this research two tests were tried neither of which were very satisfactory - the Mill-Hill, and the Wide Range. In this experiment the Mill-Hill test was used as it had been standardised in Britain - in the end only the raw scores were used.

Thus, the experiment consisted of a generation task as in the previous experiment, a rating-for-unusualness task, involving Thorndike-Lorge frequencies and Paivio's m values from his imagery norms, and the vocabulary test.

AIMS

(1) To estimate to what extent vocabulary is an important variable in word production;

(2) To determine whether original persons have a different conception of what is unusual and therefore produce more unusual ideas; - to test whether differences in generation are carried over into the subject's assessment of unusualness;

(3) To check the hypothesis, arising out of research such as Howes (1954), that subjects estimate familiarity of words with a high degree of consistency and accuracy.

SUBJECTS

42 subjects completed the experimental session of 35-45 minutes. They were tested in varying sized groups, from 1 to 8. All were students. There were 25 males and 17 females. Two subjects were subsequently rejected for failing to follow instructions, leaving 24 males and 16 females.

MATERIALS and DESIGN

Materials and Instructions appear in Appendix 4.

Blank sheets of paper were provided for the generation of unusual words. The instructions were as in experiment 1 - words had to be unfamiliar, uncommon, infrequently used in the language, they had to be 2-syllable nouns, and they had to be different from one another (to prevent chaining of words or use of plurals etc.). Each subject was asked to produce 10 words and 15 minutes were allowed for this.

Two lists of thirty words were constructed (2-syllable nouns for continuity) from 2 sources:-

1. 10 levels of frequency in the Thorndike-Lorge word count: AA, A,45, 35, 25, 15, 5, 10 in 18 million, and 4 in 18 million. These spanned the full range of values reported in the count. Three words from each level were randomly selected, and included in a random ordering of levels, constant over subjects.
2. 10 levels of m-value (associability) from the Paivio et al. count (1968) of 925 nouns: less than 3.7, 4.1, 4.6, 5.1, 5.6, 6.1, 6.6, 7.1, 7.6, and greater than 8.0. Again 3 words were selected and included in a random ordering of levels, constant over subjects.

The presentation of lists 1 and 2 were randomly arranged in the booklet supplied to subjects, some starting with list 1 some starting with list 2.
5 minutes were allowed.

Also included in the booklet was the Mill Hill vocabulary test. The test was slightly modified - only the last 24 items in each part of the scale were administered as the first 10 were judged too easy for the student population. All of these had the six choices standard in the Mill Hill test. The modification simply meant combining the multiple choice sections of Form A and Form B. All the items were presented as multiple choice questions, rather than as half definitions as in the original form. The subject's task was to underline the word which was closest to the presented word in meaning. 10-12 minutes were sufficient for this test.

PROCEDURE

Subjects, in an informal atmosphere, were informed of the nature of the task involving the use of English words (full text of the instructions appears in appendix 4).
No mention was made of time limits to the subjects, though the limits mentioned above were kept to fairly closely.
In the word generation subjects were asked to think up 10 words according to the criteria mentioned above. Subjects were stopped after 15 minutes if they had not already finished to their own satisfaction.

Each subject was then given a booklet containing the lists to be rated and the vocabulary test. They were instructed that their own ideas on unusualness were required, not what they thought the general public think. Rating was done with a ten-point rating scale with the value 10 as the least common words right through to 1 as the most frequent. They were to work through the lists quickly and not stop to ponder over words. Again there was no obvious time limit - all subjects finished within 8 minutes. The vocabulary test was then completed.

RESULTS

Raw scores appear in appendix 5. Despite the difficulties of subjects in producing words which are unusual, only one failed to produce five or more. All the responses produced were listed together, but because of the diversity of the words produced they could not be meaningfully scored in the same way as in the previous experiment - based on a count of individual frequencies (the majority of words produced would have been given a score of 3 since few were repeated). The system of scoring used was developed from the Thorndike-Lorge word count in the following way:-

a score of 0 was given to words appearing twice, or more frequent than 10 in 1 million in the Thorndike-Lorge.

1 was given to words in the range of 1-10 per million.

Correlation Matrix Table 2.1.

	Fluency	Voc.	Rat.1.	Rat.2.	OI.
Originality	.39	.18	.03	.07	-
Fluency		-.05	-.00	-.13	-
Vocabulary			-.13	-.02	.11
Rating - words chosen by m-value				.30	-.05
Rating - words chosen from Thorndike Lorge					.14
Originality Index					

a score of 2 was given to words which occurred once per million, and words not in the Thorndike-Lorge list and also not appearing in the 1944 Oxford dictionary used, but which are modern, perhaps scientific, and not necessarily infrequent.

3 was given to words which appeared 4 times per 18 million.

4 was given to words which were not in the Thorndike-Lorge list but appeared in the dictionary of the same period - this meant they were genuine English words with a frequency less than those counted in the 30000 in the Thorndike-Lorge.

A full list of words generated is in Appendix 8.
Each subject was scored for originality, as above, and for fluency. An Originality index score was calculated ((originality/fluency x 4) x 100)

For unusualness rating the measure taken was a correlation: a pearson r between each subject's ratings and the actual values (from 1 to 10) arising from the arrangement of the frequency intervals in the word counts. This can be justified since it was an imposition of an interval scaling both on the original values and on the subject's rating scale values.

These correlations ranged from .46 to .81 with the mean around .68 for the Thorndike-Lorge words. All of these are significantly different from chance. The correlation of the mean ratings of the words with the scale values given was .86, so this constitutes considerable support for the work of Howes (1954) and Attneave (1953). For the words from the m-value lists, the correlations were rather lower, ranging from .29 to .76 with a mean around .55. The correlation of the mean ratings of the words with the scale values was .63. It is clear that subjects do assess the unfamiliarity of words very accurately; indeed it is quite possible that their ratings are much more accurate than the scale values from the word counts. It is to be expected that the correlations for the m-value list are less since the scale intervals, because they are taken from a sample of only 925 words, are much smaller in subjective terms.

A fairly crude measure of the rating scale use was derived by dividing each subject's total of all his ratings by 165 (which is the expected total if the subject's use of the scale corresponds to that of the scale values of the words presented). It should thus give some indication of the subject's ability to use all the points of the scale. Correlation of the measures derived for this from the two word lists gives a figure of .62 which is significant at the .001 level. Thus the subject at least seems to use the scale in a consistent way from list to list. Although the score of 165 could arise in 2 ways (either by rating accurately throughout the scale or by rating accurately at the extremes) it would still indicate the ability to differentiate unusual and usual words. The fact that the subject does this consistently in two systems of unusualness indicates that subjectively the individual is consistent. The later test of high and low originals is more important. This test merely illustrates consistency.

The vocabulary test was simply scored as number correct for each subject.

Table 2.1 shows the initial set of correlations for all the measures. Points emerging from this:-

1. There seemed to be no significant relation between the originality measures and vocabulary - this may suggest the threshold model which was discussed above. The subjects were all students and would be expected to score fairly high in vocabulary measures, but in fact the range of scoring was from 19 (out of 48) to 46, which would seem to be adequate to indicate trends.

2. Surprisingly there was no relation between vocabulary and familiarity rating scores. Thus ability to identify the meaning of a word is separate from the ability to identify the familiarity of the word. This may be support for the coding approach to the extent that in an association model larger vocabulary would be seen in higher outdegrees for the definition of words, which corresponds to the properties of the model for assessment of familiarity.

A coding interpretation might allow that the uses of the different codes are different and therefore this result is quite in keeping with the assumptions of the model.

3. There was no relation between originality and rating abilities. On the face of it this is support for the coding approach.

4. One strange result was that there was little relation between the rating of Thorndike-Lorge words and the rating of m-value words. This may in fact be due to the fact that the scale difference, in real terms, between the m-value words was much less than the Thorndike-Lorge words, which would lead to difficulties for subjects in distinguishing points on the scale. The use of the rating scale indicates that more words were given values high in the scaling in the case of m-value words than in the case of words from the Thorndike-Lorge - perhaps a reflection of the population size 925 for m-values and 30,000 for Thorndike-Lorge. There is of course the possibility that there is no significant correlation between m-value or associability and word familiarity, though earlier research indicates that there are correlations between Thorndike-Lorge values and m-values (Underwood and Schulz, 1960).

DISCUSSION

With regard to the first aim of the experiment there is the surprising result that vocabulary, as measured by the Mill Hill vocabulary test has very little relation to originality of words generated. It is likely that this can be explained to a large extent by the nature of the population examined. A second possibility is that the measure used to determine the subject's originality is not the usual one as seen in the previous experiment and is based on norms for frequency which may no longer be applicable.

It seems that individuals producing unusual responses do not have a different conception of what is unusual from others. This was confirmed by t-tests done on separated groups of high and low originality and there were no significant differences in rating ability. It indicates that both original people and less original people judge according to the same standard; one can surmise that the problem is one of access. The original person for some reason has better or easier access to the unusual words, but when he has retrieved these words he assesses the words by the same standards as the individual with his less unusual words.

At the same time this is also support for the coding type of approach, for the reasons discussed in the preamble to this experiment. One must however temper this last remark by adding that we have only offered two models, one of which was being subjected to a fairly critical test caused by the removal of retrieval in the task presented. The fact that this model fails is taken as tentative support for the coding model, but only insofar as it should provide a more fruitful line of investigation. It has not as yet, been explicitly tested, and at this stage in time no clear predictions make it testable in this context. It is however a model for further development.

The final aim of the experiment - to investigate the status of an individual's judgement of unusualness - has more clearly been fulfilled. The correlations of subjects' judgements of uncommonness with word frequency norms, are high and significant in most cases. The mean value of 0.86 for Thorndike-Lorge and even the 0.63 mean correlation for m-values are in line with the 0.79 found by Attneave (1953) and similar results of Howes (1954). Their use of a rating scale for this purpose seems to be consistent from one situation to another. The conclusion that one can draw from this is that a subject's rating is reliable and probably as accurate and acceptable as word frequency counts in this type of research. It means that subjects are capable of retrieving this type of information directly from their memory base, and it opens up interesting possibilities for research into the fundamental individual differences in experimental stimuli such as words.

Experiment 3

This experiment was in the nature of a slight side-step from the general train of the investigation, as it was felt necessary to improve our understanding of the measures being used before launching into a fully blown examination of the coding processes. The question raised in the last experiment concerning the system of measuring the unusualness of subjects' responses is a very important one and it is clear that it must be shown to have some validity. Following from the findings of the nature of the subject's use of the rating scale to determine the unusualness of words it was decided that this could be used to indicate how successful the original scoring of unusualness had been. This required that the subjects return to rate the words produced in the last experiment.

At the same time it seemed desirable to determine to what extent subjects who produced unusual words on one occasion could do the same at a later date. In effect, could it be claimed that this is an ability, whose parameters can be reliably demonstrated on different occasions?

Further possibilities opened up: the recall of unusual words could be investigated again, with a more mature population than in Experiment 1 and further clarification of the relation of the measures to vocabulary could be ascertained by the use of an alternative vocabulary test, the Wide Range, developed by Atwell and Wells.

For these reasons the following experiment was set up.

AIMS

(1) To determine to what extent the performance in generating unusual words is a readily reproducible one;

(2) To investigate the scoring system used for the estimation of originality in the last experiment, by adopting the external criterion of subjects' own ratings;

(3) To confirm the findings of experiment 1 concerning the recall of original items;

(4) To further explore vocabulary size using a different vocabulary measure.

SUBJECTS

20 subjects from the 40 of experiment 2, chosen at random and tested alone or in small groups of 2 or 3, for an experimental session lasting 30 minutes.

MATERIALS AND DESIGN

In outline the design was very similar to the previous experiment, with a generation task, a rating task and a vocabulary test, in that order. Examples of materials are in Appendix 6.

The generation task differed in that subjects were first of all asked to reproduce their unusual words of the last experiment (up to the maximum of 10). Since this was very long term memory, recall was fairly low - though all subjects recalled at least one of the words they had produced (the mean being 3.5 words). It was then suggested that as before they should attempt to produce a total of 10 words, so that if they had recalled 3 words they should generate a further 7. This was done with a mistaken idea that it would make the generation more comparable with experiment 2. As it turned out this was not a success, since it complicated the matter of scoring, retarding the fluency of those whose memories were better and thereby reducing their potential for generation. It also created difficulties with the originality index. It would indeed be better to separate the two tasks completely, though some steps must be taken against facilitation of response generation by the initial task. Up to 15 minutes were allowed for this task, as before.

A list of 40 words selected randomly from the total list of words generated in experiment 2, was presented for rating on a 10-point scale. As before 10was least common and 1 was most common. No subject took more than 5 minutes for this section.

The Wide Range vocabulary test was then presented; it has 100 multiple choice items in much the same way as the Mill Hill, though in this instance the items are not progressively more difficult but are randomly arranged throughout the list. The test is more Americanised than the Mill Hill, and consequently certain items were consistently answered wrongly and would have to be discarded if population norms were desired. The test occupied 15 minutes.

PROCEDURE

Exact instructions appear in appendix 6. As far as possible the experimental sessions took place in the same place as in the last experiment.

Firstly, subjects were informed that the first task was as before, that they had to think up 10 unusual words, which were 2-syllable nouns and were different from one another. However, as many of these words as possible were to be the same as in the previous experiment - the subject had to recall the words he produced before. (The time interval between the two experiments was 5 weeks, a considerable period, and it was not expected that recall scores would be high. Nevertheless, it was expected that any words recalled would reflect any differential variables, such as unusualness, at work). It was stressed to subjects that though they might have difficulty in remembering the previous words they were still to make up to ten the number they produced. Average word production was around 7.5.

Secondly, subjects were asked to rate a list of 40 words on the 10 point scale of unusualness. As they were all familiar with this, little instruction was needed. It was pointed out that some of the words might be familiar as they had produced them in the previous experiment, but that they were a selection from the words that everyone had produced.

Finally, the Wide Range vocabulary test was presented and subjects simply worked through underlining the word from a group of six which corresponded most closely in meaning to a target word. 100 words were tested and the raw score was used in the analysis.

RESULTS

The main results are shown in the form of correlations in Table 3.1 (raw scores in appendix 7). The high correlation between the generated originality index and the recall originality tends to suggest that subjects do not differ markedly in their strategies, for recall, with regard to the measured variable of unusualness. This is supported by the result of a t test to gauge the differences between the expected recall originality and the actual recall originality. The expected value was calculated by the following formula: (recall fluency ÷ generated fluency (experiment 2)) x generated originality (experiment 2), which is a measure which expects words from the previous production to be recalled at random with respect to unusualness. Originality was scored as outlined in Experiment 1, in accordance with the system suggested by Cropley (1969). The matched t test produced a value of -0.3161 which is not significant. It appears that subjects do not recall more unusual words better than less unusual words, nor indeed, the opposite. Essentially, these findings are in line with those of experiment 1.

Correlation Matrix Table 3.1.

	OI 3	Wide Range	Mill Hill	Recall OI	Ratings
Originality Index T-L - Exp. 2*	.03	.11	.27	.66**	.08
Originality Index T-L - Exp. 3*		.05	.15	.13	-.07
Wide-Range Vocab.			.74**		-.01
Mill-Hill Vocab.					-.10
Originality Index of recalled words T-L*					
Rating of selected words from Exp. 2					

* based on Thorndike-Lorge word count

** significant at the 1% level.

Again there seems to be no relation with vocabulary, with the Wide Range achieving lower correlations with the measures of originality. It does have a high correlation with the Mill Hill, so that seems to be a fairly consistent measure of some aspect of vocabulary.

The problem arises however, on examination of the tiny correlation between generated originality index from experiment 2 and that produced by the same subjects in experiment 3. This would suggest that the subjects' performance was subject to uncontrolled variables in either or both of the experiments. Or in fact, that the measure taken was not indicative of any consistent ability. Related to this is the finding that the range of correlations of subject's ratings of the set of 40 words with the values assigned in the scaling used in the previous experiment, is from .28 to .65 with the mean around .45. These values are much lower than would have been expected and suggest the root of the whole problem. The measure of originality looks to be suspect, since it should conform to the ratings of subjects more closely than it does and should form a reliable indicator of performance, reproducible from one occasion to the next. It was thus decided to create a new scoring system for the analysis of unusual words, and this almost certainly must be based on the ratings of individuals.

DISCUSSION

These remarks above should be seen as a justification for the reanalysis of the data of experiment 2 and 3, in so far as it concerns the measures of originality. The use of the Thorndike-Lorge word count as a basis for the analysis has proved unsatisfactory.

It was decided that the ratings of independent judges should provide the scale values of unusualness. Independent judges were given a booklet containing 120 of the generated words (there were a total of 367 words generated in the course of the two experiments). On each page were 40 randomly selected words in a fixed order over judges, and each judge received a random permutation of three of the nine possible pages. Seven of the total number of words were not presented as they were much more commonly known as verbs than nouns, e.g. refuse, convert, compress. These were assigned the value of 2.0, as being fairly common words. Each judged the presented words on a ten-point scale, 10being the most unusual and 1 the most common. There were 15 judges and each word was rated five times. The mean of these ratings was taken as the scale value for that word. The judges were from the same population as the subjects.

Each subject's list of responses from experiments 2 and 3 were then re-scored and the appropriate test values calculated. The results are discussed in the following pages beginning with experiment 2. The re-scoring is listed in Appendix 8.

RE-ANALYSIS OF EXPERIMENT 2

Experiment 2 (contd.)

The effects of the new system of scoring can be seen in the Table presented (table 2.3). The most noticeable difference is the now significant correlation between the originality index and the Mill Hill vocabulary measure. This is to some extent what would be expected as pointed out in the previous discussion. It suggests that vocabulary should be taken into account when dealing with an experimental situation such as this.

Only the top line changes in this table from the one previously presented. It is not possible to rescore the words listed for rating, as the balance of the selection was based on the Thorndike-Lorge count, and also the purpose of the exercise was to judge the subject's ratings against these objectively established norms. The high individual correlations and the high mean (.86) indicate that this use of the levels of infrequency, taken from the whole range of values in the word count, may be meaningful despite the arguments against the more limited use of the word count to gauge the unusualness of subjects' generated words as has been found in experiment 3. It is acknowledged that there may be some grounds for using a rank order correlation of the levels with the subject's ratings, and then the product moment correlation of these values as in the final analysis table.

Correlations with Judges Values

Table 2.3.

	Vocabulary	Rat. 1	Rat. 2
Originality Index Judges	0.44**	0.13	0.30
Vocabulary (Mill-Hill)		-0.13	-.02
Rating 1 - m-value words			0.30
Rating 2 - Thorndike-Lorge words			

** significant at the .01% level.

The extent of the difference produced by this new method of scoring can be seen in the correlation between the new measure and the old one, which was only .17. It is clear that there are some very basic differences - it is hoped that this second method will prove more reliable and useful than the first. An interesting finding is that vocabulary does not affect fluency of word output. High and low vocabulary groups are not significantly different on fluency of original words.

DISCUSSION (contd.)

These results do not greatly alter the conclusions drawn previously, in that the differences in rating are still independent of the originality index and of vocabulary scores. The expected relation between originality and vocabulary has emerged and the indication is that the extent of a person's word knowledge may be important in his access to a word for generation, but may not be so important in his assessment of its unfamiliarity. This may be supported by a coding explanation whereby subjects predominantly use a semantic code for retrieval even when asked for words by their familiarity. Perhaps words are generated by meaning and then tested for familiarity, indeed it was reported by some subjects that they thought of a category and then attempted to produce specific instances of it which were unusual. What can be proposed from this is that in generating a word for no matter what purpose, one uses a semantic code for obtaining access and then tests the word according to whatever question is being asked, be it grammatical or of something more peripheral like the familiarity of the word. Thus one can confirm the outlook developed in the first two experiments that the questions of access and type of code used may be the most relevant question asked.

RE-ANALYSIS OF EXPERIMENT 3

Experiment 3 (contd.)

The re-analysed results are shown in table 3.3, some of the figures referring to experiment 2.

The first point which is apparent is that the correlation between the originality indices of Experiment 2 and 3 almost reaches significance at the 5% level (.37). With a matched t-test on these values a clear difference emerges at better than the .01 level, in the direction of higher scores in Experiment 3. To some extent this may have depressed the product moment correlation for the two values. The use of a one-tailed test of significance is quite justified here as it is a test of the reliability of the measure used and to this extent, we test a specific direction of correlation.

The relation between originality and recall originality is even higher than before indicating that there is no tendency to recall more unusual or less unusual words which had been previously generated. This is verified by a t-test value of 0.96 between the recall originality scores and an expected recall originality based on the calculation: (Recall fluency x originality index in experiment 2). The t value is not significant indicating that neither unusual or common words are being selectively recalled.

Correlations Based on Judges Ratings Table 3.3.

	OI(3)	Rec. OI	M.H.	Rat.1	Rat.2	W-R	Rat.
✱ Originality Index - experiment 2	.36	.78***	.51**	.31	.26	.50**	.25
✱ Originality Index - experiment 3		.31	.16	.41	-.01	.29	-.17
✱ Originality Index of Recalled words			.50**	.37	.09	.54**	.04
Mill Hill Vocabulary				.12	-.15	.74**	-.23
Rating of m-value words					.40*	-.16	-.03
Rating of Thorndike-Lorge words						-.18	.41
Wide Range Test							-.10

Rating (Rat) is a correlation of subjects rating with that of the mean of the judges for all 30 words

✱ scores calculated from judges' ratings

* significant at .05% level

** significant at .01% level

*** significant at .001% level

Significant correlations are found between originality indices and vocabulary measures, though surprisingly this does not apply to the originality for experiment 3. It may well be that this is in line with the lack of significance in the relation mentioned above (OI 2 VS OI 3) and in the relation with recall originality.

The usefulness of the final column in the table is not too clear as it refers to the subjects' ability to conform to the ratings of the judges. The important thing about this is that the individual agreements range in correlation from .59 to .87 with a mean around .72. Thus there is a fairly wide agreement on this measure. This "awareness" of the subjects of the ratings of others and their conformity is not related to any of the other variables measured here. The testing of unusualness is separate from access to these words.

DISCUSSION (contd.)

It may well be that the learning effect of increased originality found in experiment 3 may be due to a revision of strategy on the subject's part. The indication is that the subject is abandoning the use of simply a semantic code and using some other code which may be more successful. The work of Mednick (1962) and Torrance (1963) have indicated that there are learning effects in these type of tasks - these results may then be due to this different strategy developing. This is seen in the t-test which produced a figure of 3.15, df = 19, significant at .01 level.

Therefore there is greater originality in experiment 3 than in experiment 2.

The final conclusions to be drawn from Experiment 3 may be dealt with in terms of the aims as set out at the beginning.

Firstly it is clear that the performance in generation of unusual words is not an easily reproducible one. There is some indication that performance is consistent but this is not completely satisfactory. The implications of the differences in performance may be seen as concerning the strategies used by the subjects having changed over time, or in another way the use of a different code for retrieval. This is strengthened by the relationship which exists between originality and the vocabulary test, which is a test concerning retrieval of semantic information about a word, in experiment 2 and which does not exist between the two measures in experiment 3. The thesis proposed is then that a coding approach more easily characterises the situation found in these experiments.

Secondly, the investigation of the scoring system proved very necessary and the scoring system subsequently developed provides a more accurate description of subjects' performances than that originally used.

Thirdly, there seem to be no grounds for the notion that there is differential recall of words of differing degrees of unusualness, when these words have been thought up by the subjects. The treatment of recall in this simple manner is in all probability, too insensitive to the individual differences and to the different aspects of the coding processes inherent in the human system. The indications are that generally, in the field of memory research, studies are turning more towards these more subtle aspects of the retrieval problem, and straightforward free recall scores are becoming less and less appropriate.

Fourthly, it is not obvious what the comment here should be, since in the strict sense of the aim set out, there has been very little light shed on vocabulary size. The tests considered would seem only to deal with aspects of the semantic coding inherent in the cultural use of the word. The presentation of the word invites analysis along the lines simply of semantic feature extraction and this may not be completely indicative of the words available in the subject's word space. Indeed, work by Loewenthal (1971) confirms that there are various degrees of knowing the word even in the semantic sense, and that with particularly difficult words subjects could give an analysis based on the physical features of the word - this was seen in their errors of choice of appropriate other words. In the absence of more adequate measures of the extent of a subject's vocabulary it is realistic to use tests such as the Mill Hill, but it should be remembered that it is more appropriately a measure of coding ability of a specific nature. However, as in this experiment it can contribute to the understanding of the processes under investigation.

Experiment 4

After experiments 2 and 3 the investigation turns more and more to the problems of access and coding as the more important ones in the tasks studied. The results of the last experiment may form the basis of the assumption that semantic coding dominates, though it is possible to show that other codes of words exist and are of some value. This present experiment explores this theme in more detail.

It may at this point be worth reiterating the reasons for adopting such a task as word generation as a central part of the study. Clearly the need is to look at originality - the ability, capacity, tendency, to produce unusual material (always with the qualification that it must be useful and not simply fantastic). A definition of originality might be as proposed by Koestler (1962) - "bisociation", which is seeing an item in a different context from its usual position. The implication is that the originality is inside the person, the materials are stored somewhere, that his memory system may be the key. Ideally, a measure taken would be one which involves direct entry into the memory system.

The type of material involved in originality must be communicable; it should be capable of representation externally. This limits its presentation to modes directed at the senses - visual (pictures, design), acoustic (music), tactile (rhythm) and more generally, language systems, which require further analysis, but follow fairly strict sets of rules to make them intelligible. Because of these rule systems the symbolic mode should be easier to study, and the most refined of these and easily accessible for the majority of individuals, is verbal language.

Even with this rule system however, it is still necessary to look at simple aspects and there is scientific justification for examining the smallest units (words, phonemes, letters) as they are stored in memory. Since in this instance the production of such units must be judged for originality, in some way, the units most suitable are words - their values can be examined by the majority of individuals, who have had a great deal of experience in their use.

A final requirement is that the task be open-ended as this has been the most conspicuous base for originality, and that the subjects be allowed to produce items not normally produced in these circumstances, and hence original. A task of this nature, with the minimum of constraints designed to ensure that the subject perceives the task in a similar way to other subjects, has been presented in this research.

At this stage what is required is an attempt to elucidate what "code" means, and what individual differences are seen in individual's encodings or their retrieval using a code.

The general hypothesis is that in encoding there is some sort of precedence whereby the physical features are analysed, before the semantic importance of the item is dealt with. This is the classical view - recognisable in many of the models discussed in the review (e.g. Morton, 1968).

These ideas find some support in the work of Loewenthal (1971) who suggests and has some experimental evidence for the theory that individuals have various degrees of "knowing" a word. That is, they may not know the meaning of the word but are familiar with it to some extent, possibly can even use it correctly in a sentence. In choosing a definition of a word incorrectly there is a distinct tendency to choose one which corresponds to some of the physical aspects of the word - perhaps the sound of it, the shape of it in terms of other words which look like it, or other respects, which turn out to be indications of their idiosyncratic systems of dealing with words. This may correspond to some of the findings in this research. It certainly corresponds to the findings of Posner and his co-workers, who suggest a three-tiered system of codes - the physical, name and semantic codes. By code is meant the internal representation of a word (in this case) which is almost certainly characteristic of the way the subject perceives that word, and the way he has built up the relations of it to other words through his experience of it. These comments make the goal of research much less ambitious - rather than looking for a universal system of word categorization, which was initially a priority of memory research, we must seek only to understand the differences in the use of coding strategies which underly the memory system in different individuals.

There is then some sympathy with Morton's Logogen - where attributes are used to access the word. The more common the word, the less attributes are required to retrieve the word, the less common the word the more difficult it becomes to retrieve.

Similar too is the work reported by Brown and McNeil (1966) on the tip-of-the-tongue phenomenon, which seems to be a commonly reported state of incomplete subjective knowledge of a word. It appears when a subject, given a definition of the word is unable to recall it but can specify something about it perhaps the first or last letter. Their analysis is in terms of physical attributes being recalled but being insufficient for complete recall of the word - this is generic recall. Their discussion is not completely convincing since it suggests that this generic recall brings the definitions of words generated according to these cues to the point in memory where it may be matched to the input definition (rather like the executive decision maker of Anderson and Bower 1972). This seems to complicate the issue since to compare the definitions the features of the relevant word must be retrieved and a decision made; then presumably the word must be retrieved again for output. The nature of the representation of these comparisons is unclear; one can obtain different results if the subject has to recognise the word rather than recall, but the process suggested seems to be no different from recognition where the experimenter presents the selection of words whose definitions must be compared. However the study is of great importance since it does suggest the interaction of various functional codes devised by the subject and used according to his experience.

Since the initial experiments described above indicate that there is no difference between original persons and others in their evaluation of unusual words and in the absolute extent of semantic word knowledge, one can only assume a difference in access. If this is the case there should be differences in the use of codes for retrieval. The hypothesis adopted for this experiment is that if original persons have better access, their performance in generating when given less cues should be better than other individuals. It would seem reasonable to study this in the areas where clear coding differences have been shown previously - that is physical codes and semantic codes. In the light of research by Morton (1969) and by Conrad (1972 and 1973) it may be useful to subdivide the physical aspects up into visual and acoustic and this may be justified as indicative of the natural process whereby on encountering a word one first extracts something of the visual features of the word, then actually says the word subvocally before finally extracting the meaning. In line with the work of Loewenthal the hypothesis suggests that subjects who have better access will perform better in word generation with very basic cues like visual, than other subjects. At the same time the difference in simpler tasks like generation from semantic cues should be smaller and tend to reflect the differences in vocabulary.

The following experiment then consisted of unusual word generation, the Mill Hill vocabulary test, the generation of as many exemplars as possible of three categories, defined by their visual characteristics, by their acoustic characteristics and by their semantic characteristics. Each of these categories is balanced for size and all the words are 2-syllable nouns. A final task was to estimate the size of the subject's set of these words so that he was asked to indicate the extent of his knowledge of a selection of the categories. In this way a realistic measure of availability of the words would be obtained.

AIM

The aim of the experiment was to investigate differences in the use of codes in retrieval of words from an almost completely specifiable set. Some clarification was sought of the existence of a relation between the production of unusual words and words produced according to various postulated encoding attributes.

SUBJECTS

34 subjects completed the 2 parts of the experiment; of these 6 were rejected for not following instructions. The analysis was thus conducted on 28 subjects. These were all students, the majority being 1st year Psychologists at the beginning of their course.

MATERIALS AND DESIGN

Materials and instructions are shown in Appendix 9. The experiment consisted of 2 separate sessions of 30 minutes, completed in most cases with a gap of 3-4 weeks.

The first session was in line with the previous experiments - there were 2 tasks: firstly generation of unusual words, subject to the usual constraints that the words should have 2 syllables and be nouns; secondly the Mill Hill vocabulary scale.

In the second session the subjects were tested individually and the first task was to generate 3 sets of nouns. The categories from which these words were taken were previously investigated in a pilot study. In this, the purpose was to establish the generatability of different specifiable sets, corresponding to three forms of coding: visual, acoustic and semantic. By sampling from the Thorndike-Lorge word count of 30,000 words, six sets of roughly equal size were established and these were tested for subject acceptability. There were 2 from each of the codes:

visual - words of 7 letters in length,
words with 2 letters, going below the line and none appearing above, i.e. words with the letters g,j,p,q,y but not b,d,f,h,k,l,t - e.g. pig, grip, quay, here it was explained to the subjects that this would give them an idea of the shape of the word, and this shape was pointed out.

acoustic- words which have 'gh' or 'ph' sounding as 'f' but not appearing at the beginning of the word, as this would make the task one of articulation rather acoustic decoding (this being the most normal type of word experience for the individual),
words which have 'ch' sounding as a soft sound, as in birch, as distinct from hard as in chemistry (again this sound was not allowed to appear at the beginning of the word).

semantic- words which are a type of container
words which signify an emotion.

In the case of visual codes 2 letters below the line was chosen since it allowed subjects to generate the word directly from the shape code. With the length of the word subjects seemed to be examining words randomly till they found one of the correct length by counting the letters. With acoustic, the 'ch' sound was chosen as it gave rise to less problems being clearly defined in English; the 'gh' in contrast caused a degree of trouble since it may have a different harder sound in Scottish or Irish pronunciation, as in "sough" and

"Lough Neagh", "ph" had too many Greek roots. The semantic words caused little trouble and the containers were chosen as being closest to the others in set size.

The actual set sizes was discovered by a laborious count of the instances found in the Thorndike-Lorge count for the visual and acoustic codes, while the semantic set size was determined by cross referencing in Roget's Thesaurus. The sizes found in this manner were around 160 instances, which in the course of the experiment went up to around 180, by dint of other acceptable words which the subjects produced. The words in each category are listed in Appendix 12.

It was hoped that the difficulty of the generation tasks with these categories would be equalised by using total sets of the same size. The total set sizes found were visual - 190 words, acoustic - 182 words and semantic - 172 words. When these were broken down into different frequencies of occurrence, X^2 values show no significant differences between the categories.

For each category whose presentation order was randomized, subjects were allowed $4\frac{1}{2}$ minutes to write down as many words as possible which conformed to the instructions given. The time was determined from the pilot study where subjects' production and interest began to wane after $3\frac{1}{2}$ minutes. The adjustment of time limits is necessary in the light of Guilford's (1971) finding that using long time limits simply gives a measure of vocabulary size rather than fluency or, in this case, generation according to the codes provided.

A measure of vocabulary size was obtained from the next task which consisted of presentation of 30 words from each category randomly ordered, beside which each subject indicated his knowledge of the word. He did this by putting the letter A, B or C beside the word, for the following reasons:

A - he knew the word and could define it;

B - the word was familiar but he was not sure of the meaning;

C - he did not know the meaning and had never seen the word before.

The use of this strategy to investigate the subject's word knowledge was reported by Loewenthal (1971) and her results suggest that the subject quite accurately indicates the extent of his knowledge, and the scheme is meaningful to subjects. The reasons for using this technique here were firstly, since subjects would have different particular vocabularies it is important to ensure that each subject knew the majority of words in the categories (any subject who deviated markedly from the average words known could be rejected from the sample) and secondly, it allowed an estimate of the total number of words from which a particular subject could choose his responses (rather than using the total obtained from the Thorndike-Lorge etc.).

Further to the experimental subjects, another 27 from the same population were asked to be independent judges of the unusualness of the words generated by the experimental subjects in the first session. The judges were given no indication of where the words had come from, though it was stressed that they were not under test in any way. The rationale for using the ratings of independent judges has been discussed in the previous experiment. Each judge rated 160 words on a 10-point scale of unusualness. The words were presented in a booklet of 4 pages of 40 words, the pages being randomly arranged for a specific judge. Each word produced by the experimental subjects was rated 8 times, at least.

PROCEDURE

The instructions given to the subjects are shown verbatim in appendix 9. The first session was similar to the tasks described in the previous experiment, involving the subjects in thinking up as many two-syllable nouns as they could which they considered unusual. They were allowed 10 minutes for this task - output varied from one or two words in the time to over 30, with an average around 12. The Mill-Hill vocabulary test was then given with the reservations previously mentioned in mind; subjects were allowed to finish the test, though none extended the testing session beyond 30 minutes.

The second session conducted around 3 weeks later involved the subjects individually - initially it was hoped to analyse the latencies of word production via tape recoruings of the subject's voicing of the words as he produced them, but the data collected proved too complex to be easily analysed in the time available. Some comments on the results of this aspect of the experiment can be seen in Appendix 18, where some attempt is made to fit the general shape of the data into the work of Johnson and Johnson (1951) and Kaplan and Carvellas (1969) and also Broadbent (1973).

It was explained to the subject that he had to produce words according to certain instructions that he would be given, there would be three types of words and he was to produce as many as possible in the time till he was told to stop, after approximately 5 minutes. No examples of the categories were given to the subject though the sound necessary for the acoustic category was demonstrated, and words which were disallowed, like chemistry, were pointed out. The letters involved in the visual category were mentioned and it was stressed that this should give him some picture of what the word should look like. The words produced were written down by the subject after he had said it aloud.

Finally a booklet of three pages of words in block capitals, drawn from the three categories, was presented. In this the subject had to specify his word knowledge according to the scheme discussed in the previous section. The words were selected randomly from the total lists available.

Judges were enlisted to rate the words produced in the first session, most were students but came from a variety of disciplines and were asked to rate according to how familiar they were personally with the words in the lists. Each judge was given around 160 words to rate - no difficulties were encountered in this. The mean ratings were used to score the experimental subjects' productions.

RESULTS

The raw scores appear in Appendix 10. The analysis takes the form of a series of analyses of variance, shown in tables 4.1-6. Correlation of all the main variables is in table 4.7.

The analysis was done on the results of the total number of words generated by each subject when given each code instruction, and on the words known by the subject in each of the categories. The analysis used is a split-plot with the vocabulary scores and the originality index scores being formed into groups at the quartile points, to give 4 groups of 7 subjects each.

Table 4.1 and 4.2 investigate the effect of grouping according to the tests taken in the first session and should partially answer the question posed in the introduction to this experiment concerning the relation of generation of unusual words (a task with very little constraint) and generation according to certain constraining variables postulated to influence our encoding of words.

Analyses of Variance - split-plot design

(1) Words Generated   Table 4.1.

Source	df	Mean Square	F	
Vocabulary (4 levels)	3	128.13	2.52	n.s.
Error (a)	24	50.71		
Code	2	1045.36	57.73	< 1%
Code x Vocabulary	6	13.01		n.s.
Error (b)	48	18.10		

(2) Words Generated   Table 4.2.

Source	df	Mean Square	F	
Originality Index	3	193.40	4.54	< 5%
Error	24	42.55		
Code	2	1045.36	65.30	< 1%
Code x O.I.	6	29.81	1.86	n.s. < 10%
Error	48	16.00		

(3) Words known   Table 4.3.

Source	df	Mean Square	F		
Vocabulary	3	818.19	2.78	n.s.	7%
Error	24	293.86			
Code	2	2841.33	28.74	< 1%	
Code x Vocabulary	6	185.85	1.88	n.s.	
Error	48	98.83			

With vocabulary as an independent variable there was no effect of vocabulary on word generation - a larger vocabulary does not mean that more words will be produced when category information is provided. It probably suggests that indicators designed round knowledge of random words have no predictive validity for production of words from specific categories. This is to some extent what would have been expected in the light of the discussion in the previous experiment. The effect of different code information is clearly significant, with the semantic code producing the largest number of responses, followed by acoustic.

Table 4.2 shows that originality index scores predict word output in the fixed category situation, with higher originality having greater word output. This effect is significant at the 5% level. The effect is not due to the common base of fluency of ideas, since the correlations, shown in Table 4.7, between fluency, "measured in the unconstrained word generation task", and the code generation scores are not approaching significance. The implication is that the more original the subject the more easily he can deal with retrieval by code, or theoretically, that there is a common factor of ability to retrieve when presented with code information. Again the effect of the codes themselves is significant at well beyond the 1% level. There is no interaction effect.

Tables 4.3 and 4.4 take words known by the subject, measured by his A responses to the list of words presented in session 2, as the dependant variable. The number of A responses is used to estimate the total number of words known in the particular category. The measure is (Words Known ÷ 30 or 31) x Category Size. The values 30 or 31 arise because the word "pitcher" included in the rating list could belong to either containers or words with "ch". In the case of containers there were thus 31 exemplars presented. Table 4.3 indicates some effect of the general Mill-Hill measure on the responses to the word list presented, though the relation is not significant. On examining the relevant correlations, it is clear that the lack of relation of vocabulary with the knowledge of words in the acoustic code, may have made general word knowledge across the codes unaffected by groupings according to the vocabulary measure. Again the code variable produces a large effect, though surprisingly this is due to the high level of knowledge of the acoustic category. More will be said about this.

Table 4.4 indicates that ability to produce unusual words does not affect the extent of the subject's word knowledge.

Tables 4.5 and 4.6 repeat the findings for the words the subject has previously encountered, in the categories. However, the code effect is now attributable to greater knowledge of both the visual and the acoustic, or conversely to the lack of knowledge of the semantic code words.

Analyses of Variance

Table 4.4.

(4) Words Known - taken from A responses

	df	F	
Originality Index	3	0.75	n.s.
Error	24		
Code	2	26.86	< 1%
Code x O.I.	6	1.25	n.s.
Error	48		

Table 4.5.

(5) Words Known - taken from A + B responses

	df	F	
Vocabulary	3	2.84	< 10%
Error	24		
Code	2	30.16	< 1%
Code x Vocabulary	6	0.70	n.s.
Error	48		

Table 4.6.

(6) Words Known - taken from A + B responses

	df	F	
Originality Index	3	0.18	n.s.
Error	24		
Code	2	31.81	< 1%
Code x O.I.	6	1.15	n.s.
Error	48		

The relevant figures relating to the category size, knowledge and production are in Table 4.8.

Table 4.7 presents the results in the same manner as in previous experiments. Several interesting points emerge:

(1) The surprisingly small relation between words known in each category and its corresponding word production score. These 2 measures are not significantly related in the acoustic code and only at the 5% level on a one-tailed test for visual and semantic codes.

(2) The relation between Originality and word generation is apparent in the significant correlations with both the acoustic and semantic generations. However the visual category does not fit into this pattern.

(3) Fluency in unusual word generation does not relate to words produced in the categories. This is strange in that they are both measures of the number of words produced in a fixed time in response to a set of instructions implying a specific set of words. However in the light of Guilford and Hoepfner (1971) report of a factorial study on fluency, it could be argued that it is to be expected - they show that in a comparable task, when there are no restrictions in symbolic cueing, word fluency loads only .22, whereas when constraints increase (in line with those here) the loading jumps to around .70.

Correction Matrix Table 4.7.

	Fl.	Vocab.	WG Vis.	WG Ac.	WG Sem.	WK Vis.	WK Ac.	WK Sem.
Originality Index	-.41*	.28	.28	.63**	.38*	.32	.29	.19
Fluency		-.06	.13	-.14	.05	.35	-.11	.19
Vocabulary			.36*	.50**	.50**	.54**	.21	.45**
Words Generated - Visual				.58**	.44*	.35*	.02	.14
Words Generated - Acoustic					.43*	.35	.20	.17
Words Generated - Semantic						.63**	-.04	.36*
Words Known - Visual							.36*	.63**
Words Known - Acoustic								.24
Words Known - Semantic								

* significant at 5% level

** significant at 1% level

(4) Corroborating the above there is a clear cluster of correlation within the word generation factors - one can assume that these are more clearly indicating the presence of word fluency, in the form found by other researchers.

(5) Mill-Hill vocabulary score seems to relate to the majority of the measures concerning the codes. While to a certain extent this is understandable with the words known, it seems to conflict with our previous results concerning the words generated.

(6) There seems to be no reason for the very significant correlation between words generated according to semantic codes and words known visually.

Table 4.8 probably provides the most interesting results of all. From this it would appear that the size of the set from which the subject generates is not directly related to the magnitude of his actual output. That is, comparing words known (average over subjects) with words generated in each of the categories the differences are clear. In terms of the initial aim of the experiment subjects generate words more easily when they are given semantic information about the words required, than when given acoustic and this better than visual.

It is also apparent that the size of the set may be determined in different ways according to the criterion of word knowledge and there may be some grounds for supposing that words may be known in a great many degrees, whose values in terms of familiarity, may form a continuum.

Important too is the fourth column of Table 4.8. Here it can be seen that there are more different words produced in reponse to acoustic cues than to semantic or to visual. This might suggest that acoustic information may allow access to a larger area of memory despite the fact that it does not provide suitable cues for production.

Table 4.8.

Summary of Words Generated and Known

	Total words possible	Average words known	Average words generated	Different words	Percentage different words
Visual	190	148	8.00	91	45
Acoustic	182	163.57	13.21	134	35
Semantic	172	144.71	20.18	115	21

DISCUSSION

The point emerging from the experiment is that production of words is easier when a semantic code is given for the word, than when an acoustic code is given and that the provision of only visual cue for words is of comparatively little value. The general implication is what might have been expected, that there is some support for the notion that the memory system is most usefully accessed in semantic codes. This is essentially in line with the models of Morton (1969) and many others, and there is clear evidence that the use of the semantic code supercedes set size effects in determining output.

This result is essentially what was looked for in that it seems intuitively appealing that the hypothesised course of word encoding (visual analysis, acoustic analysis, and then semantic encoding) should correspond to the ease of recall when codes providing information at each of these levels alone, are given to the subject.

The Originality index scores seems to serve as a good indicator of general output level in word generation, indicating a better ability to retrieve when given code information for a limited set, is given. It is interesting that no interaction effect appears in Table 4.2, thus there is no selectively better performance on any one of the codes by those who are original. It was suspected that there may have been a difference in the more difficult code situations where access to more unusual items may have been a help. One surprising aspect of this result is that the above relation is not due to fluency scores. One may have tended to ascribe the relation between these two types of measures of open-ended tasks to the total number of responses produced in these situations. However, there is no significant correlation between the fluency score which is part of the originality index and the measures of word generation. As suggested before this is probably due to the fact that the word generation tasks can be identified with word fluency in the Guilford sense, while the lack of constraints on unusual word generation makes it not a fluency test at all. The fact that the word generation measures are highly intercorrelated tends to confirm this hypothesis. Thus it seems that there may be a general factor of use of coding strategies for retrieval which underlies these tasks, the extent of use perhaps providing the index which is available in the production task.

One final problem remains and this concerns the unusual relation between words generated and words known in the three code situations. That is, though the order of magnitude of production was semantic, acoustic then visual, the words known or the total set size from which the subject generated was, acoustic, visual then semantic. If we simply ascribe this to the fact that the semantic encoding is better we are sidestepping the issues, since this tells us nothing about the organization in memory. The codes refer to words which are present in long term storage and are accessible, - there is no clear reason why one code should produce better results than another, unless the structure of the address is different in the case of semantic. No one can doubt that the semantic is the information which the subject uses most in everyday life and to some extent he must be more familiar with its use. However, we come back to the fact that the words being output according to other codes are just as familiar - why then is there greater output when there is a smaller number of words to choose from? Further concern about this can be expressed when the figures for the number of different words output by the subjects is examined. Interestingly when these figures are translated into percentages of the number of words produced in each category, it would appear that visual cues produced the greatest diversity of response suggesting that the code information did not provide access to the same area of memory in different individuals.

The type of model proposed here to account for these facts is a distance one. Essentially, it proposes that when words are learned over the course of a number of years they tend to be located in memory storage close to items of similar usefulness, or in most cases, similar meaning. Thus memory clusters are formed round semantic usefulness and confusions etc. common in all memory tasks are due to the close proximity, and close meaning of the words dealt with. The full implications of such a model will be dealt with in the final discussion. Here it requires only to say that it would neatly account for the data found and that from it there can be derived certain testable predictions.

According to the work of Sternberg (1966, '67) a useful method of testing notions of distance in human memory, is to examine reaction times. This would mean that the reaction time to identify a word given either visual, acoustic or semantic information about it will vary as the distance apart of the various categories, and not as one might have expected according to the order of extraction of the various codes in temporal order (visual, acoustic then semantic). This prediction will be examined in the following experiments.

One interesting cross-reference is to the data shown in appendix 18. In one of the final graphs the number of words in units of time over the whole course of the word production, is shown. Each category appears separately. What is important is that though the semantic word production begins at a much faster rate than the other two by the end of the time allowed, there seems to be no difference in the rate of production. In fact this is the case from approximately 1 minute 40 seconds onwards. This would agree with the model proposed in the sense that words which are close together in "space" will be produced quickly and easily at the beginning of the search, but as the search continues and a general area is exhausted, the words will have to be chosen from a larger area even more randomly. The effect of fatigue etc., would not explain the disproportionate decrease in output from semantic codings. Nor is there an answer in the fact that the semantic category is being reduced in size by the subject's output, since the average overall production is only 20 words when the estimated set size is over 140.

A further suggestion that the most frequent words have been exhausted and only infrequent words remain, making the task much harder, may be dismissed by examination of the order of production of the words by the subjects and the frequent words which subjects omitted.

The distance model being proposed then, is similar to Meyer's (1972) spreading wave of excitation in that concentration on a specific area initially produces more results with a semantic code but as this area is passed over leaving common exemplars, the advantage of the semantic code disappears, and one would expect to see variables of set size coming into play, as the search becomes more random. This is in fact the case, and acoustic output becomes slightly better than semantic, over the last minute of the production time.

Experiment 5

From experiment 4 the need for a model to explain the strange results, is obvious. What I wish to present is a distance model for 2 reasons:

(a) because of the lack of correlation between measures of category size and no. of words generated.

(b) the fact that the number of different words produced is far greater in acoustic than semantic - even though subjects were producing more individually in semantic, they were selecting these from a smaller subjective set. A neat way of explaining this is to postulate the existence of clusters of words in the memory base which are probably arranged semantically. So in the case of say acoustic encodings, the words are equally available (to semantic encodings) but more widely dispersed. Thus it is likely that the subject in searching for the appropriate words, samples from various clusters, which would lead to a greater divergence of output words.

In the final analysis what is left is a model which will include the process of search (and the individual differences therein) and which will also give a method of explaining the above points.

The proposed model consists of 2 main parts: firstly a memory base from which all very well known items (in this case words) are retrieved and secondly a search process which is a repetitive probability-based mechanism.

The Memory Base This will contain all the words used by the S, either as word units or as some other basic phonemic unit. (It is also possible that the representation consists not in the contents of an address - in computer terms - but in the address itself. This type of conceptualisation would find support with Learning Theory models such as Rosenblatt (1958), and Physiological theories such as Pribram (1972)). The structure of the base will be determined by 2 things:

1. the distance between units,
2. the strength or familiarity of the unit.

Initially the distances may be randomly allotted to the emerging units, but increasingly the ordering of things will be determined by association and in the end by familiarity and frequency of association. This will lead to an arrangement of the units which will rely on the 2 basic concepts, above, and which should also be highly efficient in reflecting the norms of word occurrence in every-day language.

The actual "spatial" arrangement should be in terms of usefulness and use which in most cases will lead to an arrangement by meaning - giving rise to the associative clusters extensively investigated by Palermo and Jenkins (1967). We can now deal with the research on associative clusters, and things like the "spew" hypothesis of Underwood and Schulz (1960) since our model will allow their findings but also help to explain the fact that the first word out in an associative spew is not the same in all cases of the same S. Our model says that the position of entry into the network determines the distances to an item and therefore the probability of its output.

Intuitively one would expect the type of arrangement to be dependant on the coding of the input and this itself will be time-based to allow for all the effects found in the distinction between STM and LTM (Baddeley and others 1966). It then seems sensible to propose that there is a sequential analysis involving processes of visual, acoustic and semantic encoding, the most fundamental being semantic but at least some of the effect of the physical encodings being present in the final analysis in the memory base. This would give us a flexible entry into the base.

The problem of the nature of the retrieval from the base, in terms of the processes existing beyond the base (c.f. Kiss's omission of processes into and out of his thesaurus) must be left for later discussion.

<u>The Search Mechanism</u> This derives from ideas gained in the analysis of experiment 4 - the output of words falling into the categories appeared in groups of 3 or 4 in much the same way which Broadbent (1973) suggested; and the cumulative record showed the characteristic curve reported by Kaplan and Carvellas (1969) and others. In simple terms it would seem that the subject stumbles on a group of appropriate words at certain points in his search and these are output. One would expect these to be related and though

this has not been tested statistically, it would appear to be the case. One could imagine that the search drops off in intensity as new words are not found and therefore we reach an asymptote which is a great deal less than our capacity. But this is a speculation.

A second idea which appeared was that the Original Subjects tended also to be the ones who produced the most output, (this is in fact a very common finding in studies of Originality and Mednick (1962) found that the final responses in the subject's listing were the ones most likely to be original). This final point relating to Mednick would tend to support the theory of the memory base proposed above, in that the least likely responses would be produced later in the output and these would be the furthest away from the first output point. To the extent that Originality ties in quite closely with motivation (studies by Barron 1963) the idea of increased search (in time) as a function of increased motivation is appealing.

This is very helpful, since the problem is in explaining the very large individual differences found in output of words whether these are to fulfil the function of an experiment or whether it is a literary venture - since each individual subject has a very similar word store both in size and in structure and has a very similar accurate system of judging frequencies of occurrences and literary merit. That is to say, differences in production of words is not to be explained in terms of storage or even in terms of encoding, but rather in degree of search undertaken.

The nature of the search process is in a simple excitation of the pathways or addresses in the location of the starting point, and spreading for as long or as far as the strength of the impulse allows. At each address or pathway a threshold decision is made as to whether to output an item or not. The decision is based on the two factors in the memory base. The degree of search is determined by the number of times search is initiated when looking for an item.

The aim of experiments 5 and 6 is to test the ideas involved in the memory base concept set out above. Experiment 5 initially looks at the recognition time for words when the words have to be identified by verbalization, and experiment 6 concerns recognition of the categories (codes) without necessarily identifying the particular word involved. Thus experiment 5 looks at types of coding in the memory base while experiment 6 deals with the coding processes prior to entry into the memory.

METHOD

Subjects: 10 Ss were run in a control group and 10 in the experimental group. Each of the 10 experimental Ss had taken part in experiment 4, so scores were available for them for word generation etc. All Ss were students at Bedford College.

Design: 3 categories of word were presented, corresponding to the codes of the previous experiment. They were:

visual - 2 letters below the line, none above;
acoustic - 'ch' sound, but not at the beginning;
semantic - a type of container.

20 words were presented from each of the above categories, selected from the set of all words generated in the last experiment, in each category. The 20 words were selected as follows:
experimental group - as far as possible 10 words* were taken randomly from the words generated by that particular S in the previous experiment for each category. A further 10 words were selected at random from the set of all words generated in that category in the last experiment. (When more than 10 words were not available in the subject's previous responses, all words generated were included, and the number made up to 10, if necessary, with words randomly selected from the total set. These words were then treated as further control words.
control group - 20 words were selected at random from each code set.

For both groups, words were arranged in groups of 5 i.e. 5 cards, one word per card, and a counterbalanced, random ordering presented in 2 halves of the experiment, separated by 2 minutes i.e. visual (5 words), semantic (5 words), acoustic (5 words), acoustic (5 words),

* some subjects did not generate as many as 10.

semantic (5 words), visual (5 words) - 2 minutes break then a different random ordering. In each group of 5 words the "generated" and non-generated words were randomised. An example of this is shown in Appendix 13.

There were 2 reasons for this grouping:

1. a brief pilot study indicated that Ss required a number of instances of the category together in order to adopt a meaningful strategy (if the words were completely randomly arranged they would not reach an optimal strategy because of the constant switching).

2. it was also important that the groups were not too large, since the task is in the end fairly monotonous - it would therefore be useful to break up the task (see vigilance research for confirmation). Also if the ordering had been random, constant interruption to state the expected category, would make the experiment too long. The counterbalanced design was, as usual, to control for fatigue effects.

In the experimental group subjects were informed which category of words were to appear in the next group of 5; the control group were given no indication of grouping or of the categories involved. Although the main purpose was to compare the 3 category effects, it was important that the effects were due to the instructions given rather than to any inherent differences in the particular words used. Thus the control were to allow the comparison of no fixed strategy, with knowledge of the codes to be expected.

All subjects received 15 practice trials, to allow RT to settle down and to allow an estimate of threshold of recognition to be made. All words presented in the experiment appeared for 0.3 secs., the practice having established in the majority of cases that words could be reported correctly at exposures of 0.03 secs. Thus all presentations were well above threshold.

Apparatus: A 3-field tachistoscope made by Electronic Developments Ltd, though only 2 fields were used, with an automatic card change. Words were on cards of the appropriate type and thickness and were printed in lower case lettering - Letraset 14 pt. Helvetica Light. Sizes are about $\frac{1}{4}$ inch. All words appeared in the centre of the visual field.

Recognition times were measured by a voice key interruption tuned to the subject's speech; the tachistoscope impulse started, and the voice key stopped, the centisecond digital timer.

Procedure: This followed the experimental design closely. Each subject was given a unique selection of words, randomised as above, to take into account the words he had previously generated. Several (6) of the control subjects received the same words in the same order as experimental subjects, i.e. the subject was matched to a control subject. This arrangement would have been ideal for all control Ss but the difficulty of obtaining subjects at fixed times (in advance) made the time investment in this operation too great to be contemplated.

All subjects were questioned at the end of the experiment whether they had guessed the purpose of the experiment (this was especially important fcr the control subjects, who were given no indication of the grouping of the different types of words) and in all cases they had little knowledge of the purpose.

Results

Raw data is shown in Appendix 14. Analysis was carried out to examine differences in recognition time for generated and non-generated words and also to compare coding differences. The predicted results were faster recognition for generated, and faster recognition for semantic over acoustic over visual.

The first prediction suggests that generated words will tend to have been familiar words and thus will be recognised quickly when re-presented. The second prediction is derived from the assumptions in the model of distance, to the effect that the semantic encodings would be more closely arranged in "word space" and therefore knowledge of the area in which to search would lead to faster output in the case of the more closely related words i.e. semantic - the other ordering follows from the idea of precedence of encoding and the results of earlier experiments. See the description of the model for the reasoning in detail.

Analysis took the form of repeated-measures analysis of variance as shown in Tables 5.1 - 3.

Table 5.1. Experimental group only

Source	df.	MS.	F	
Subjects	9	627.02		
Code (sem,ac,vis)	2	144.81	8.95	< 1%
Code x Subjects	18	16.16		
Generated vs Non-Gen.	1	365.06	9.62	< 2.5%
Gen. etc x Subjects	9	37.91		
Code x Gen.	2	30.21	2.34	n.s.
Code x Gen. x Subjects	18	12.29		

Table 1 gives the analysis carried out on the experimental group. The effect of having generated the words previously is significant at the 2.5% level; if we are to accept the "spew" hypothesis, this effect is likely to be the result of some more basic variable relating to the subject's lexicon, but it can be taken as an illustration of the importance of taking into account the particular structures and weightings predominating in the subject's set of words.

As predicted, we can see a significant effect of code with semantic faster than acoustic in turn faster than visual (see graph 1 for confirmation). Using Scheffe's test for unplanned comparisons (1953), which is a very conservative test, we obtain the following as comparisons on the treatment totals:

Table 5.1A.

Comparison	D_i	MS.	
Vis. vs Acoustic	32	51.20	n.s.
Vis. vs Semantic	105	551.25	< 5%
Ac. vs Semantic	73	266.45	< 5%

5% level mean square = 114.74

Thus the difference is attributable to the difference between the semantic and the physical codes in general, rather than our distinction between acoustic and visual. Further comparisons on codes, having separated generated and non-generated words (i.e. on examination of components of the Code x Generation interaction), do not reach significance.

Table 5.2 shows the analysis on the combined data from the experimental and control groups. A point to note here, is that, since the random selection of words were for the control group, a mixture of generatable and non-generatable words (i.e. no sample of their generation of the words was taken, so one can only speculate on the extent of the words which they would have generated given the chance), the scores used in this analysis for the experimental group, were weighted individual means (balanced for generated and non-generated totals). These values can be seen in the appendix. The analysis is a split-plot.

Table 5.2. Analysis for Experimental and Control groups

Source	df.	MS.	F	
Exp. vs Control	1	317.40	1.53	n.s.
Error	18	206.43		
Code (vis,ac,sem.)	2	84.46	9.65	< 0.5%
Code x Exp. etc.	2	29.40	3.36	< 5%
Error	36	8.74		

Thus, we see no overall difference between the experimental group and the control group. This suggests that the cues given to the experimental subjects are not particularly helpful in aiding recognition - in fact the overall tendency is for the control subjects' recognition times to be faster. However, since there are extremely large variations between subjects in recognition times, the result may not be as important as comparisons within a subject.

Comparisons on the treatment totals, comparing the experimental with the control group on the 3 codes separately, were carried out - results are in the form of 'F' values for the comparisons and appear at the top of graph 5.2. The F values are less than one might imagine because the design of the initial analysis was a split plot while for the comparisons the error mean square had to be taken as treatments x subjects which is much larger than in the split plot analysis (since one now loses the variable of subjects). None of the comparisons is significant using a planned comparisons rationale.

It is important to consider why it could be more difficult to identify the word when the code was given. The result seems to suggest that the code did not prove a useful key to the particular word, i.e. that the word was not encoded in that way in the memory base. This is perhaps to be expected in the memory base or lexicon and the results (figure 5.2) indicate that the smallest difference between control and experimental groups is in the semantic category.

This discovery is that semantic cues come closest to the particular classification system used by the individual and that acoustic and visual are less important, in that order. The overall difference between experimental and control groups probably arises because the process being tested is normally automatic (identification of incoming words) and that making the process overt and reportable, increases the time for identification.

The code differences appear as significant. But the significant interaction and figure 5.2 tend to suggest that only the experimental group has shown a significant effect. This is verified by the separate analyses shown in Table 5.3. Partitioning the interaction, results in a significant effect of semantic vs visual + acoustic, but no effect in visual vs acoustic. Thus the effects can be attributed to a "physical" code difficulty.

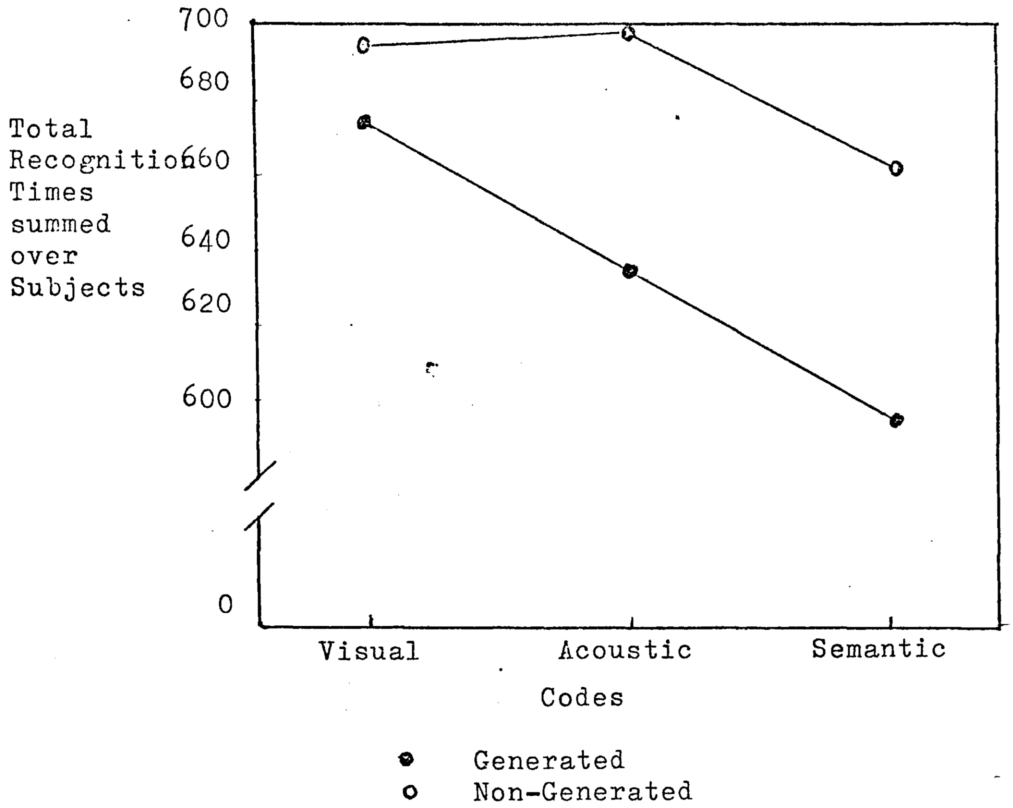


figure 5.1. Total recognition times for the experimental group, for coded words.

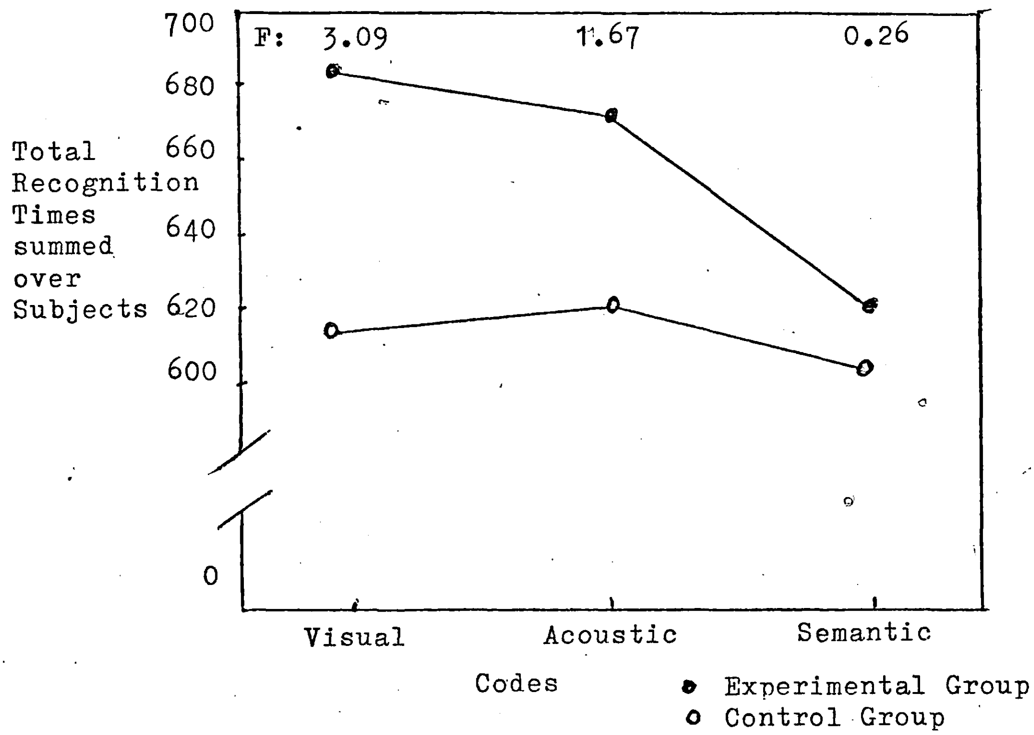


figure 5.2. Total recognition times for control and experimental groups for combined generated and non-generated words.

Table 5.3. - separate analyses

Source	df.	MS.	F	
(a) experimental group				
Subjects	9	271.03		
Code	2	104.43	10.41	< 0.5%
Code x Subjects	18	10.02		
(b) control group				
Subjects	9	141.83		
Code	2	9.43	1.26	n.s.
Code x Subjects	18	7.47		

Thus no significant effect is found in control subjects.

Discussion: In summary then, there is no evidence to suggest that announcing the coding strategy improves overall recognition time, but it does serve to create an effect for each different category. There is a significant difference in recognition time between physical and semantic conditions in the experimental group. Before considering whether this result upholds the proposed model we should examine 2 common explanations for this type of phenomenon:

(a) category size

(b) word frequency.

(a) Category size is an internal indicator of the number of words available for output in a certain context. The estimate used in this instance is derived from the subject's responses (A,B or C) to the list of randomly selected words from the total possible categories. The actual estimate was calculated by taking a proportion of the words known

(A responses) over total possible, for each of the frequency groupings and adding. Range of the scores assigned to the subjects is shown in the appendix. The category size approach gives an almost perfect fit to the "non-generated" data in graph 5.1. The explanation would be of the form - the smaller the category through which the subject has to search, the faster would be the recognition time. It would seem to be clear that when the subject does not know the word i.e. it is not familiar immediately (and this may be the key point) he searches randomly through the words available till he finds the correct one. Here the category size explanation may be useful.(1) However, it does not account for the data for "generated" words in graph 5.1. There is also a problem in accounting for the control group results in graph 5.2. In spite of this, it was decided to look at the implications of this a little further. From the previous experiment estimate of category size could be obtained and used as a covariate in analysis.

(1)In effect it looks as if category size is only useful when the subject recognises that the words to be identified fall into a specific grouping, and when the needs of the task (e.g. time limits) or the unfamiliarity of the items dictate that a random search is necessary. In some studies this would be characterised by increased errors, but here because of the long presentation interval it is seen in the RT differences.

Thus category size may not be an explanation of the RT effects but simply the outcome of a particular strategy used - in this instance - random search.

(b) In order to test whether word frequency effects could be an explanation, the following X^2 tests were carried out on the Thorndike-Lorge frequencies and the number of words which fell into each frequency from the various groups presented to the subject.

Table 5.4 shows the result. From this it can be seen that there may be cause for further investigation of the effect of word frequency - since there is a significant difference in the nature of the words. Word frequency effects reflect the results of the comparisons on the treatment sums. Thus the faster recognition times to acoustic and semantic may be predicted if we accept that the more common the word the faster it will be recognised (a normal finding in research on this topic) - it does account for the part of the graph, acoustic-semantic being parallel for generated and non-generated.

But the non-significance of the final X^2 shows that word frequency may not explain the main effects of the experiment.

Table 5.4 - X^2 for the numbers falling into the Thorndike-Lorge frequency (AA,A, 26-50, 11-25, 1-10, less than 1)

Comparison	X^2	Significance
Vis.Gen.[d] vs. Vis.Non-Gen.[d]	9.398	n.s.
Ac. Gen. vs. Ac. Non-Gen.	20.366	1%
Sem.Gen. vs. Sem.Non-Gen.	20.637	1%
Generated words		
Vis - Ac - Sem	14.641	n.s.
Non-Generated words		
Vis - Ac - Sem	14.361	n.s.

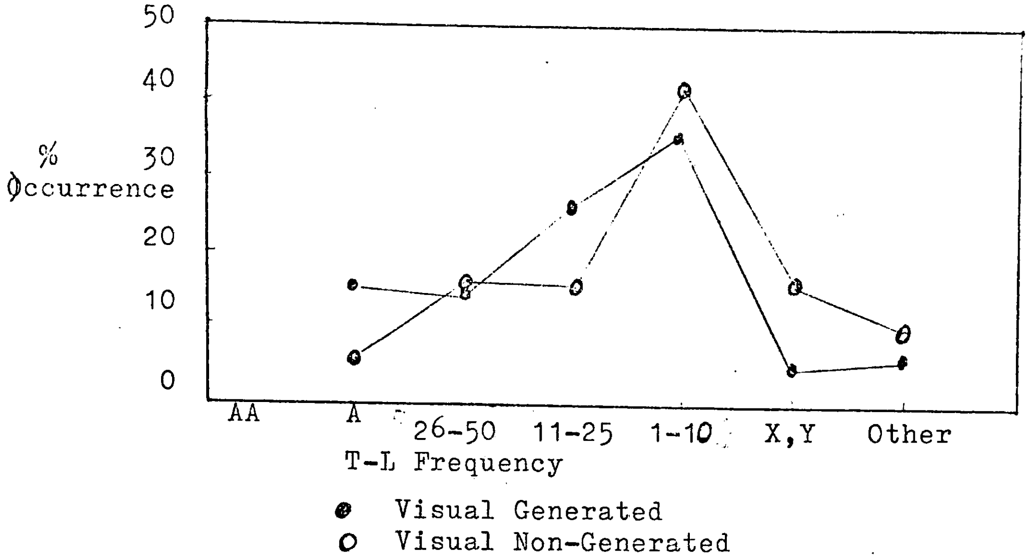


figure 5.3. Test of Word Frequency
Visual generated vs Visual non-generated

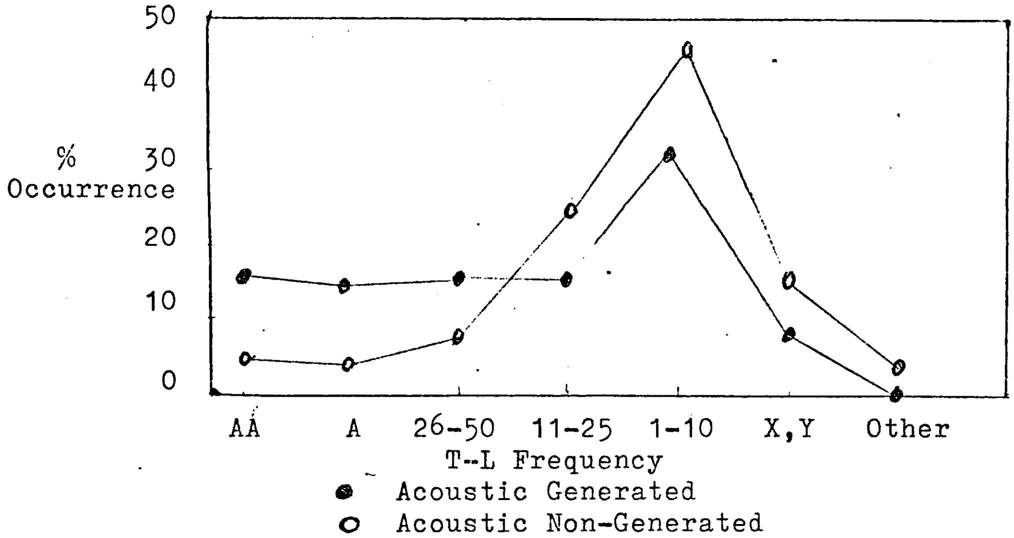


figure 5.4. Test of Word Frequency
Acoustic generated vs Acoustic non-generated

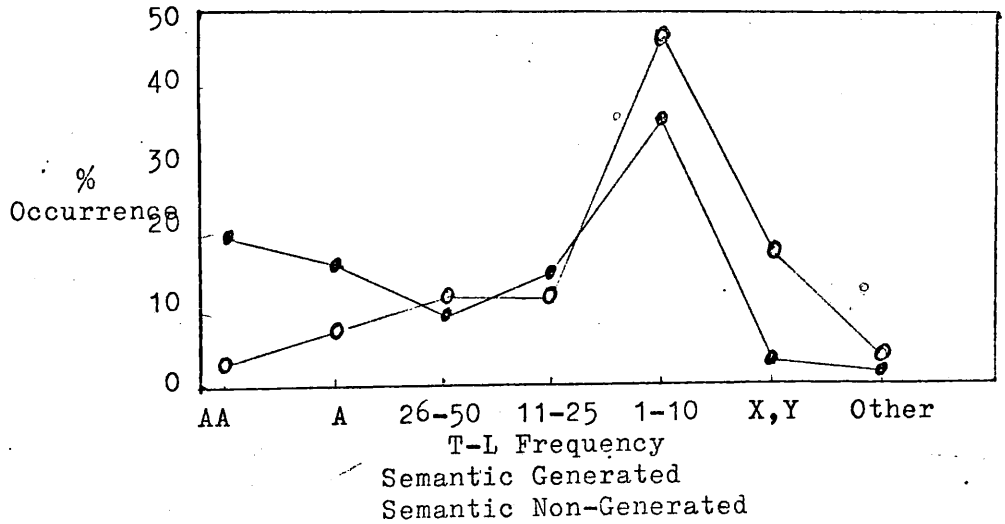


figure 5.5. Test of Word Frequency
Semantic generated vs Semantic non-generated

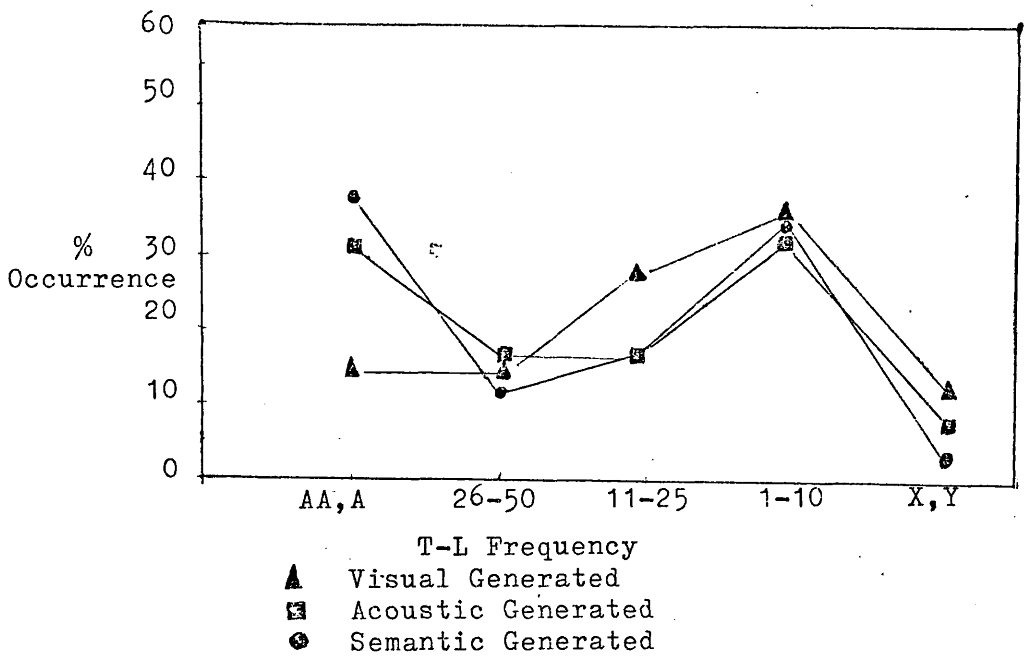


figure 5.6. Test of Word Frequency
Generated Words

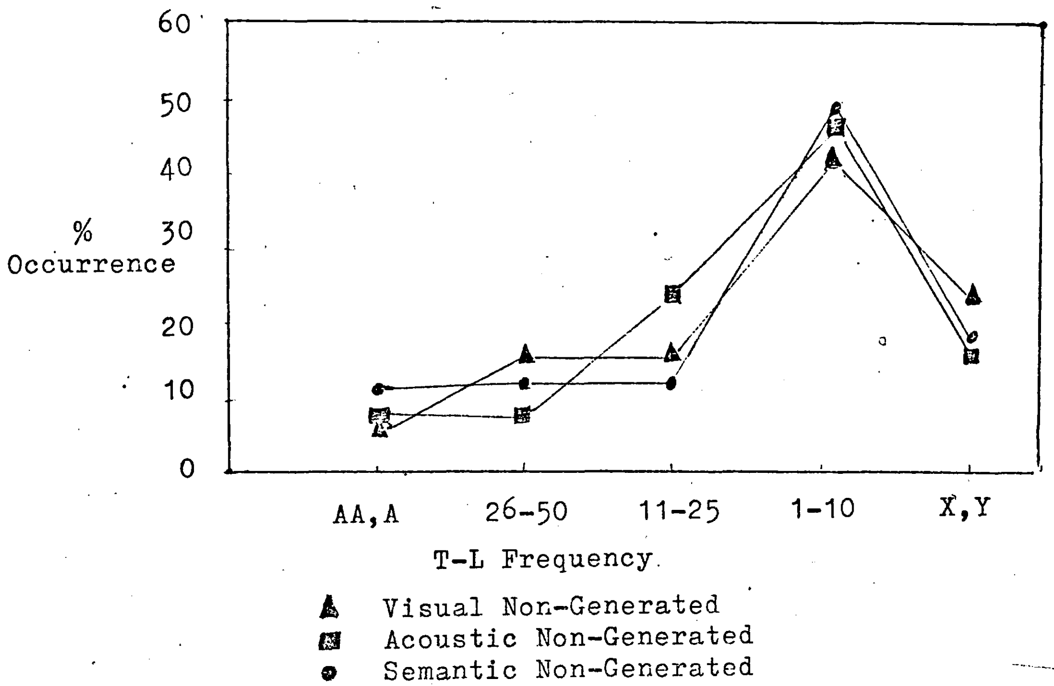


figure 5.7. Test of Word Frequency
Non-generated Words

Graphs of the patterns tested above are shown in figures 5.3 - 5.7. From figure 5.4 and 5.5 one can see the pattern of frequencies, where generated has more high frequencies and less low frequencies. Overall then, one would expect faster recognition.

As with category size, the results are far from clear-cut; therefore it was decided to break the initial data down into the frequencies and carry out further analyses to try to gauge the extent of the effect of word frequency.

To this end and also to take into account the category size effects, analyses of covariance were carried out.

Results

Table 5.5 shows the new analysis of variance taking into account the frequency divisions. Because this analysis has been done as an extra investigation, there are a number of missing data points in the analysis; in order to make the results meaningful and also to retain the repeated-measures design, a number of frequency divisions have been collapsed. The analysis has been done with 3 frequency divisions - 26 and over, 11-25 and less than 10 (all these are values from the Thorndike-Lorge G count). Even with 3 divisions, 11 data points were estimated* thus losing 11 degrees of freedom. However these were 11 points in an analysis of 180 and according to Cochran and Cox (1957), the analysis should not be greatly affected by the inaccuracies of this.

*The method of estimation of missing data points is attributable to Yates (1933).

Table 5.5. Analysis of Variance: of recognition times, for categories broken into frequency of occurrence.

Source	SS	df	MS	F	
Subjects	14489.8	9	1609.9778		
Generation	273 .8	1	273.8	7.6533	<.005
Generation x S	321 .97	9	35.77		
Code	426 .84	2	213.42	4.3167	< .05
Code x S	889 .93	18	49.44		
Frequency	1150 .87	2	575.43	19.9484	<.005
Frequency x S	519 .23	18	28.84		
Generation x Code	132 .4	2	66.2	1.0131	n.s.
Gen x Code x S	1176 .15	18	65.34		
Gen x Freq.	.83	2	.41	.0054	n.s.
Gen x Freq x S	1387 .72	18	77.09		
Freq x Code	175 .15	4	43.78	.7433	n.s.
Freq x Code x S	2120 .73	36	58.90		
Freq x Code x Gen.	399 .86	4	99.96	.8781	n.s.
Freq x Code x Gen x S.	2846 .24	25	113.84		

Table 5.5 shows the initial analysis for this. From it we can see quite clearly the effect of word frequency, significant at the 0.5% level. The code effect is somewhat reduced but is still contributing to the results, and the effect of generated vs. non-generated words again is significant.

Table 5.6. Analysis of Covariance: one covariate - set size of subject for each frequency

Source	SS	df	MS	F	
Subjects	14445.66	9	1605.07		
Generation	273.8	1	273.8	7.6533	< .025
Gen x S	321.97	9	35.77		
Code	302.4	2	151.2	3.1973	n.s.
Code x S	851.20	18	47.28		
Freq.	660.66	2	330.33	11.4599	< .005
Freq x S	518.84	18	28.82		
Gen x Code	132.4	2	66.2	1.0131	n.s.
Gen x Code x S	1176.15	18	65.34		
Gen x Freq	.83	2	.41	.0054	n.s.
Gen x Freq x S	1387.72	18	77.09		
Code x Freq	175.03	4	43.75	.7461	n.s.
Code x Freq x S	2111.52	36	58.65		
Gen x Code x Freq	399.86	4	99.96	.8429	n.s.
Gen x Code x Freq x S.	2846.24	24	118.59		

Table 5.6 shows the analysis of covariance with the set size of each subject for each frequency as the covariate. The set size variate has no effect on the size of the generated etc. effect since there is no difference in set size for generated compared to non-generated. The importance of the covariance analysis is now seen in the fact that the code effect does not now reach significance - it seems that when one extracts the variable of the size of the set from which the subject must choose his response, the differences

in the coding strategies disappear.. Thus, one can clearly adopt the category size explanation, together with word frequencies as accounting for the effects discovered in this experiment.

The magnitude of the word frequency effect is reduced also, suggesting some correlation between it and the covariate. No interaction effects approach significance.

Discussion

To summarise the findings: having taken out the effects of word frequency and the size of the set from which the subject has to choose his response, we find the following effects survive:

1. the words which the subject has already generated are recognised significantly more quickly than the words which are new to the subject.
2. the addition of the covariate, set size, serves to remove the general effect of code but comparisons on the adjusted code totals reveals that there is a significant difference between identification of the word using the visual code and the other codes (visual taking longer) (Table 5.7).
3. as a general trend the use by the subjects of the codes tends to hinder recognition, but only the difference for visual proves significant between experimental and control groups.

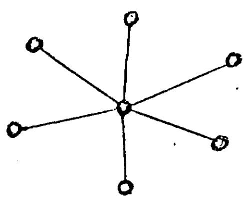
Table 5.7. Comparisons on adjusted Treatment Totals

	Visual	Acoustic	Semantic	F	Signif.
Adjusted totals	3972.67	3792.67	3828.67		
Comparison	+1	0	−1		
	$D_i = 144$	$D_i^2 = 20736$	M.S. = 172.8	3.654	< 5%
Comparison	+1	−1	0		
	$D_i = 180$	$D_i^2 = 32400$	M.S. = 270	5.709	< 5%
Comparison	0	+1	−1		
	$D_i = 36$	$D_i^2 = 1296$	M.S. = 10.8	0.228	n.s.
Comparison	+1	$-\frac{1}{2}$	$-\frac{1}{2}$		
	$D_i = 162$	$D_i^2 = 26244$	M.S. = 293.5	6.22	< 5%

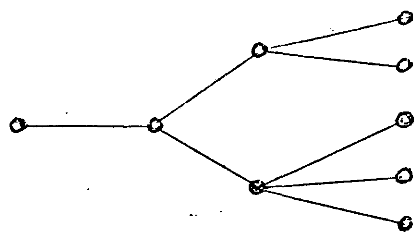
4. a significant effect of word frequency is found thus bringing our findings into line with previous research (Broadbent 1972)
5. a final analysis on the percentage of words generated from the available set gives a significant effect of code i.e. generation is more efficient from semantic codings than from acoustic ... visual.

Looking at these results in greater detail in terms of the theory which we wish to propose, we see that there is support for a distance explanation coupled with the idea of familiarity. Point 1 illustrates the idea of subjective familiarity, as distinct from word frequency and shows that we are dealing with a variable which is probably unique w.r.t. each individual. See the statement of the model for a complete specification of the nature of the "familiarity".

As regards the effect of codes, graph 5.1 gives the picture - our choice of interpretation at the moment lies between a category size or a distance model. It looks as if the 2 ideas are confounded in that as the category size increases so the distance between items increases - this seems likely though it is by no means essential. For example the difference between the 2 arrangements:-



Size increasing without distance



category size increasing with distance

Having agreed that the category size (C.S.) interpretation is more useful in recognition than in generation (see before), let us further narrow down the point of application. Since the CS approach does not account for the results of the control group, I think it is fair to say that the CS is only useful to the subject when he sees that the words to be identified fall into a specific grouping and therefore a category search becomes a useful strategy. Further, on the basis of the difference between generated and non-generated, it is important to add that only when the unfamiliarity of the items dictates that a random search through the "possibles" will be fruitful, does the CS consideration come into being. In fact the usefulness of the CS approach is dependant on the strategy used by the subject in a particular instance. One other important point here, has been mentioned (see note 5 above) - the usefulness of the code is determined by its correlation with the subjective coding system. This correlation seems to be greatest with the semantic and least with the visual - as a result when we take into account the CS as in the covariance analysis the major effect is on the semantic and the acoustic, reducing a previously significant effect (see table 5.1a) to non-significance. The remaining difference as in point 2 above, is between visual and the rest, thus it looks as if the CS interpretation is useful only when the category is also a subjectively useful one. The fact that CS goes some of the way to disentangling the code effect illustrates the complexity of the matter.

As far as the distance model goes, it would seem to give a neater explanation of the results. Looking at figure 5.1 again, and comparing the generated to the non-generated, one can argue that since the words are less familiar, then the distance from the core concept is further, therefore the RTs are longer (simply a broad explanation of the obvious results). More precisely, one would expect bigger differences in RT between semantic generated and semantic non-generated than between acoustic... than between visual...

Since the visual words are widely spaced in the lexicon the differences between familiar to core concept, and unfamiliar to core concept will be small relatively whereas with semantic the distance difference will be relatively much larger. This is in fact what is found - the differences are in the order semantic (greatest) then acoustic then visual. Though the difference between semantic and acoustic is not great.

In summary, what is argued is that the need for taking into account the individual differences in the lexicon is most important. Secondly, that the recognition of words according to specific category cues is best interpreted in terms of the distances in word space and the subjective familiarity of the items rather than the category size and the objective word frequencies.

Point 5 simply upholds the idea that distance is likely to be a better predictor of word generation than category size and in fact is the basic consideration in the internal system.

Experiment 6

Consistent with the notion of the two systems of search and the memory base is the analysis of incoming material which is prior to search - and therefore distinct from the concepts of distance. This stage has been variously identified in iconic stores,precategorical acoustic storage and implicitly in the code system proposed by Posner. The distinction is essentially between the processes acting on the input in preparation for its use in the memory system and the effects of the organization in the memory system, which are seen in the nature of the output produced.

This experiment arose out of the need to obtain more information concerning the precise nature of the effect which could be attributed to the presence of the three types of code dealt with in the previous experiments. It is important to decide whether the codes themselves (and therefore the initial encoding, or analysis, of the input) have an inherent value as organisers in memory and aid output, or whether the effect can be attributed to the action of further more basic variables which relate to the memory base (such as distance, familiarity etc.). One way of looking at this is to examine tasks which do not necessarily involve entry into the memory base and note the effects of using the three codes. This was done in this experiment by using a task which involved the extraction of the coding features from, but not necessarily the identification of, a word.

One point which should be stressed is that the task did not necessarily require the subject to use the memory base in the normal manner of word identification - with the semantic category the processes involved must be close to the memory base ones if we admit that memory is arranged predominantly in semantic fashion. However, it is quite in keeping with the earlier formulations that decisions may be made about the coding links in memory without output of the word structures and therefore without the effects inherent in the organisational system. It is important that the claim here is only that the subject need not enter the memory base to make his decision.

The experiment took the form of each subject generating words from the 3 categories as had previously been done and then in a fixed 2-choice situation, indicating whether briefly presented words belonged to the stated category. If the claim could be made that the tasks had no relation whatsoever to the memory base, the model would predict that the difference in latency for category membership would follow the same course as the analysis - i.e. visual would be fastest, then acoustic then semantic. However, because of the weaker claim discussed above the prediction of the model is adopted as no difference in latency for identification of category membership. Thus one would expect that there would be no effect of word frequency, of set size or of ability to generate words in each of the categories.

Aim: To establish the nature of the codes used for generation and their relation to organisation.

METHOD

Subjects 15 subjects randomly selected from a pool provided for exam candidates. Subjects were paid to be available for a 6 hr. period. Testing took place during this period. All Ss were good (probably native) speakers of English.

Design Each subject was tested individually in a 45-minute session. This was divided into 3 parts.
(a) an initial period of generation according to the 3 usual categories described in experiment 5 (visual, acoustic, semantic). For each category, given in random order across subjects, 3 minutes were allowed. This was rather less than in previous experiments simply because of time constraints. However it may be noted that correlations between 3 and $4\frac{1}{2}$ minute time limits in experiment 5 were .92, .95 and .92 for visual, acoustic and semantic respectively. This part was included to obtain further information about the processes in the output of words and in order to investigate any relationship between generation and the task of identification.

(b) Word presentation. 15 words from each of the generated categories listed from the previous experiments and appearing in appendix 15, were randomly selected from 6 frequency intervals as follows:

T-L G count	AA	A	26-50	11-25	1-10	X
no. of words chosen	3	2	2	2	3	3

There was no set rationale behind this distribution, except that it should be the same across the 3 groups and that a full frequency range seemed desirable. Also there were larger numbers of exemplars in the low frequency groups. For each word a non-category member was randomly selected from the T-L count and matched for frequency. The complete list is shown in the appendix. In all 90 words were presented to each subject.

Using the same rationale as in experiment 5, random selections of category and non-category words, in groups of 5 were presented in a random counterbalanced design of "abccba". For each subject there was a 2-minute break after every 30 words. Across the 3 parts the design was adjusted so that each category occupied the "a", "b" and "c" positions only once.

Each subject received the same 90 words but the 3 parts were randomized across Ss. Latencies of the subject's decision of whether the word belonged to the specified category were measured with a voice key.

(c) The final section entailed the standard measure of category size - 90 words 30 from each category, beside each of which the S had to write A, B or C according to whether he could define the word or whether he had seen the word before or whether he had never encountered the word.

Apparatus As in experiment 5, the words were presented on an automatic card change tachistoscope at an exposure of 0.3 secs. The tachistoscope impulse started the centi-second timer and the circuit was broken by the subject's yes or no via the voice key.

All words were presented in the centre of the card - the print was Helvetica Light 14 pt, and all the words were in lower case lettering.

Procedure Each subject received the tasks in the order set out in Design. The subject was introduced to the experiment: (see Appendix 15)

"This experiment is in 2 main parts - firstly you will think up some words and secondly I will show you some words and ask you to decide whether they belong to certain categories."

The normal instructions for the word generation followed as detailed in experiment 5. Three minutes was substituted as the stated time limit for generation. In the second section the subject spoke his response into a microphone which acted as a voice key and stopped the timer. The instructions were as follows:

"This part involves you in looking at some words, and telling me whether they belong to these categories which you have just dealt with.

This machine is a Tachistoscope and what it does is to flash up cards for a brief period of time. Now you will have time to read what's on the card but it won't appear for a long time.

On each card will be one word and what you have to do is to say whether it belongs to a specific category. The cards are arranged in groups of five (consecutively) and before each group I will say which category you have to look for. So your answer will be to say YES or NO, yes it belongs to, say, the category of containers, or no it doesn't.

Now, this thing here is a microphone which stops a timer as soon as you say yes or no, and what I am interested in is HOW LONG it takes you to decide whether the word belongs to that category. So say as quickly as you can whether the word belongs to that category.

The procedure will be that I will say right, and the word will appear in the next couple of seconds."

There were no specific instructions given about errors - time to recognise was the main variable mentioned to the subject. Error rates varied from subject to subject but overall were low. Any queries about the apparatus were dealt with. In the final part the category size estimator was given to find out "how many words you know as you cannot be expected to think up words if you have never seen them before."

RESULTS

The analysis used was similar to the last experiment and was a repeated measures analysis of variance. 3 degrees of freedom were lost because of missing data points caused by errors in collecting the data - mainly due to seemingly random errors with the voice key i.e. a stray mains impulse occasionally stopped the timer at the beginning of a presentation.

The original data is shown in appendix 16. The dependant variable was decision time in centi-seconds, for the yes decisions. The analysis of variance was carried out on 15 subjects, but since one S exceeded the limit of errors on word generation his scores on this could not be included in further analysis. Thus the final tables present analysis on 14 subjects. The results appear in Table 6.1 - 6.5.

Table 6.1. Analysis of Variance for all experimental variables: done on RT.
(Because of missing data 3 df are lost)

Source	SS	df	MS	F	
Frequency	1364.81	5	272.96	1.4532	n.s.
Freq x S	13148.57	70	187.83		
Code	882.25	2	441.12	.6211	n.s.
Code x S	19885.63	28	710.20		
Freq x Code	6774.14	10	677.41	4.4025	< 1%
Freq x Code x S	21080.62	137	153.87		
Subjects	39126.49	14	2794.74		

Table 6.1 shows the result of the initial analysis. As predicted by the model the coding condition is not significant. This indicates that there are no significant differences in the time to verify category membership or in terms of the model, time to analyse according to a predetermined strategy. Given these points it supports the idea that one does not have to enter the memory store in order to test whether a word has certain characteristics. This is verified by the fact that there is no effect of word frequency in this task. This is clear support since information concerning familiarity of a word can only be held in store, since it is not inherent in the word but depends on the pattern of usage.* Thus the subject is not obtaining all the information available on that word, but more likely, is simply analysing for the specific feature requested for that word.

One point here is that the analysis has been carried out only on the "yes" data. Derivations for the negative decisions would require some estimation of a stopping rule for analysis of the presented word, since this is crucial to the decision time. This may vary for each category, and in the absence of some method of estimation the treatment of these results is beyond the scope of this analysis.

* This may only be a first approximation as Landauer and Freedman (1973) indicate that infrequent words may have different structure.

One puzzling feature of Table 6.1 is the appearance of a significant interaction. On examination of Figure 6.1 one can locate the cause of the effect around the latencies for the first three acoustic frequencies. Thus acoustic codes are faster for high frequency words, but slower for A words and words appearing between 26 and 50 in the frequency count. The main part of the effect would seem to be the latencies for the latter group. There seems to be no reason for this. The words used were: "scratch" and "discharge" - one would not have thought of as being out of the ordinary. And examination of their presentation position gives no further clue. The latencies are slightly longer to "discharge" than to "scratch" but this is to be expected as "discharge" appears at the beginning of a group of 5. Using Scheffe's test for unplanned comparisons the difference between the 26-50 frequency group for acoustic versus visual and semantic, is significant at the 5% level, but neither of the other frequency differences is.

Very interesting here is the close similarity between the latencies for semantic and for visual. Classically they would be ascribed to different levels of processing - one being inherent in secondary memory and the other being a physical analysis usually linked to primary memory. This will be discussed further.

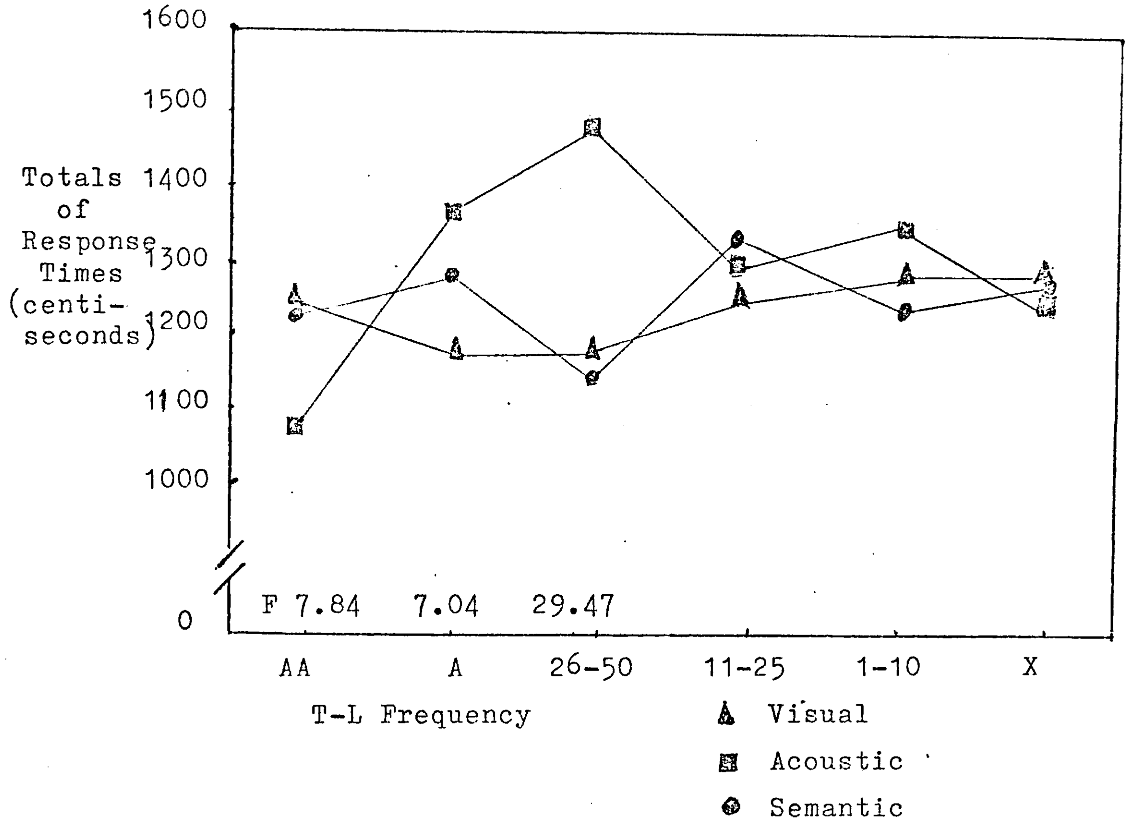


Figure 6.1. Totals of subjects decision times according to word frequency and code

Table 6.2. Analysis of covariance: where covariate is estimated set size of the subject for each frequency range

Source	SS	df	MS	F	
Subjects	38358.61	13	2950.66		
Frequency	1867.32	5	373.46	2.0167	n.s.
Freq x S	12037.10	65	185.18		
Code	678.75	2	339.37	.4568	n.s.
Code x S	19313.99	26	742.84		
Freq x Code	5919.31	10	591.93	3.9513	< 1%
Freq x Code x S	18875.60	126	149.30		

Table 6.2 shows the results of the analysis of covariance suggested by the findings of experiment 5. Tables 6.1 and 6.2 are not completely comparable, because there are only 14 subjects in the covariance analysis. The outcome however, does not alter the view of the main effects: frequency is not significant and neither is code. This would be expected following the reasoning above - if the task does not deal with the processes of the memory base, the actual size of the categories from which the subject can search for his response is not relevant to the problem in hand. Thus it transpires that the covariate of set size does not add anything to the analysis.

Tables 6.3 and 6.4 illustrate the effects of further measures from the memory base. Neither result alters the picture seen in the first analysis - the verification of class membership does not depend on the variables which have been associated with the memory base. Table 6.3 gives the data for a covariance analysis using the number of words generated as the covariate. As it has been hypothesised that this measure concerns how an individual can retrieve a word from the lexicon, it would not be expected to have a marked effect. Our suppositions are supported. Table 6.4 illustrates the combined effect of the 2 covariates is very small and indeed the interaction effect is larger than when set size is the only covariate.

Table 6.3. Analysis of covariance: covariate is no. of words generated in each frequency

Source	SS	df	MS	F	
Subjects	38697.53	13	2976.73		
Frequency	1814.19	5	362.83	1.9864	n.s.
Freq x S	11872.94	65	182.66		
Code	1096.69	2	548.34	.7453	n.s.
Code x S	19128.72	26	735.72		
Freq x Code	6004.78	10	600.47	3.9946	< 1%
Freq x Code x S	18940.79	126	150.32		

Table 6.4. Analysis of covariance: 2 covariates a) set size b) no. of words generated

Source	SS	df	MS	F	
Subjects	38314.35	13	2947.25		
Frequency	1442.89	5	288.57	1.5617	n.s.
Freq x S	12011.20	65	184.78		
Code	748.33	2	374.16	0.5063	n.s.
Code x S	19214.20	26	739.00		
Code x Freq	5999.65	10	599.96	3.9926	< 1%
Code x Freq x S	18783.44	125	150.26		

Table 6.5. Analysis of variance of words generated: scores analysed are divided by appropriate set size for frequency and category. Thus scores are percentage generated of possible words in S's set

Source	SS	df	MS	F	
Subjects	604.46	13	46.49		
Frequency	9423.44	5	1884.68	46.88	$< 1\%$
Frequency x S	2612.60	65	40.19		
Code	5900.72	2	2950.36	65.49	$< 1\%$
Code x S	1171.16	26	45.04		
Freq x Code	3864.42	10	386.44	10.41	$< 1\%$
Freq x Code x S	4821.69	130	37.08		

Table 6.5 casts some light on the nature of the subject's output in word generation. The figures analysed are the percentage of words available for generation in each frequency interval, which were actually generated by the subject. Firstly, Ss generate a significantly higher percentage of semantic coded words than acoustic than visual - as might have been suspected by the previous results. It would appear that search for words based on a semantic cue is more efficient. Secondly, Ss generate a significantly higher percentage of words available in the more frequent groups than in the less common groups. Again this is not startling, though one should emphasize that it is not the same as the findings of the "spew" hypothesis - it is information confirming the importance

of familiarity on the subject's word production and gives a more accurate indication of the subject's efficiency in doing this. Finally, an unexpected interaction appears. This can be examined through figure 6.2 - it would appear to be caused by the high proportion of semantic words which are produced in the high frequencies. This is mainly seen in the AA category, otherwise the curves are almost parallel. This would be predicted by a model which took both familiarity and distance into consideration, in measuring the ease with which words are produced.

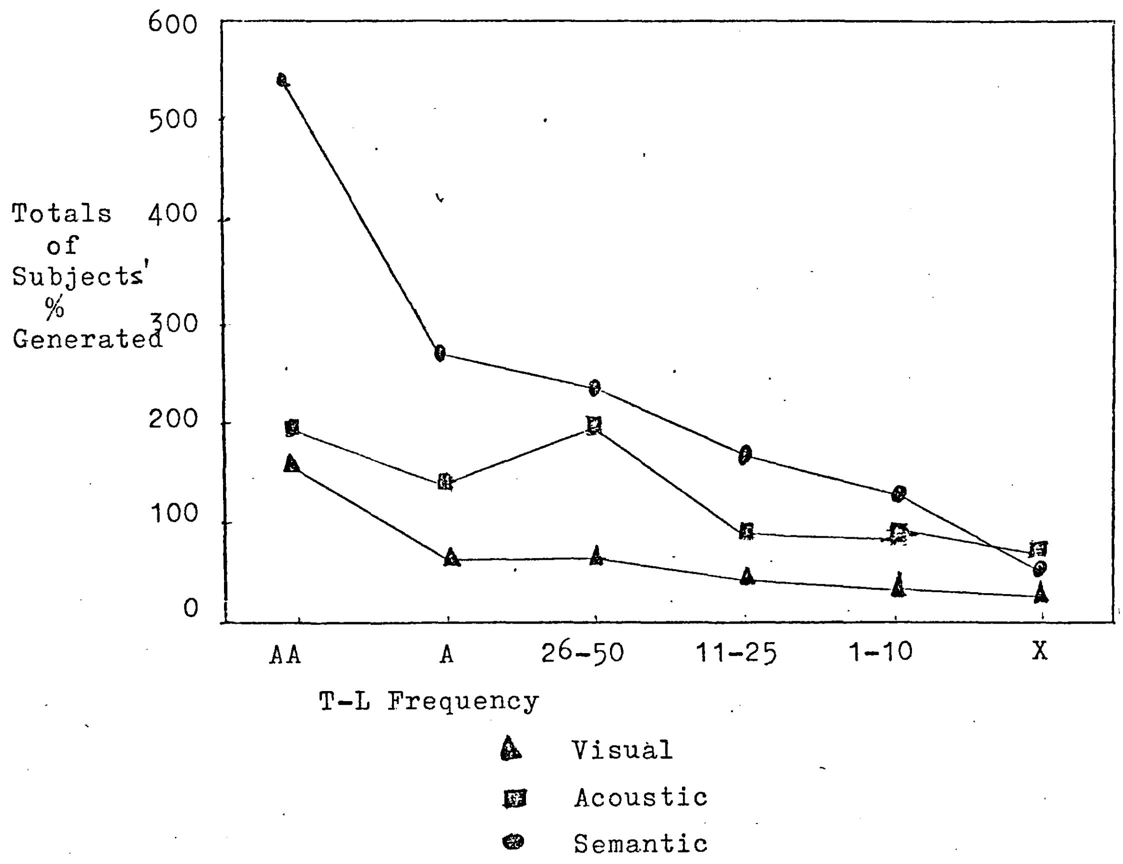


figure 6.2. Totals of percentage of words known which were generated.

DISCUSSION

Quite clearly the results of the experiment place the nature of the task outside the processes normally ascribed to the organizational properties of the memory system. In a certain sense it is strange that the work of Meyer (1972) is not duplicated in the findings concerning the effect of set size on decision time. The nature of the task he used was slightly different in that what was presented was a logical statement which may have implied for the subject a necessary test of category membership involving testing of category exemplars which may require entry into the lexicon. His explanation of his results in terms of set size may then be quite acceptable.

However, the measure of set size used in this experiment is probably closer to the actual figure for each subject, and therefore, it is deemed acceptable to reject the influence of category size in the production of the results. What this means is that the task need not involve entry into the lexicon, and therefore it is proposed to identify the task as one which involves only the analysis of the input word, in terms of the particular characteristics suggested to the subject.

Bearing this in mind, the results would then imply that the subject can turn his abilities of analysis equally to each of the particular attributes inherent in the stimulus. To some extent this conflicts with the initial idea that there was some sort of precedence in the encoding of a stimulus word - i.e. first, visual then acoustic, then, semantic. However it may well be that these are processes which have become automatic and largely subconscious. The verbalization of such tests may increase the decision time markedly (this would apply more to visual than to semantic, which is a more realistic task in everyday terms). One other possibility, is that the verbalization time is far greater than the analysis time and therefore hides the differences in the different analyses, assuming the 2 processes can overlap.

The value of this experiment then, is in the fact that it tends to place the effects of the different categories found in experiment 5, firmly in the area of the memory base. This would support the notion that memory is not simply arranged in terms of semantics but may also hold information concerning the physical characteristics of the words. If it is then concluded that the main variables involved in the processes tested in these last 2 experiments, are not the codes themselves or something inherent in them deriving from the encoding process, the question arises as to what these might be.

Perhaps the final analysis held in Table 6.5 has the answer. Here it can be seen that when the number of words available to the subject is taken into consideration at several levels of objective word frequency, an effect of code is found. If this is not due to the codes themselves as suggested by the analysis on the latencies for decision, and cannot be dealt with by the notion of set size, further constructs must be used. These further constructs it is held, are distance and familiarity which directly affect word output in the manner described in the statement of the model. This whole conceptualisation will be discussed in the following section.

GENERAL DISCUSSION

It is clear that the latter part of the study has produced the more easily interpretable results, and that from it can be devised a testable model of the basis of memory. At the same time the latter part has some important contributions to make to the understanding of the initial three experiments. With this in mind the Discussion set out here will examine first the implications of the last three experiments.

To recap on the findings table 4.8 is presented again. This was the initial trigger to this line of thought leading to the postulation of the distance model.

Table 4.8.

Code	Set size	Words Generated	No. of different words generated	% age of total words produced which were different
Visual	148	8.00	91	45
Acoustic	163.57	13.21	134	35
Semantic	144.71	20.18	115	21

As the ability to produce words in the three categories is related significantly (Table 4.7) it would appear that the same process is being used in all cases. The interesting points are then that the semantic information provides a better cue for word production than the others, and this effect supercedes the effect of set size. Simply to say, as has been done by Psychologists, that semantic is naturally better as a cue, is not particularly useful in this instance. "Why should it be better?" is a more realistic question since in the context presented the semantic cue is simply code information about stored materials, and since in using codes the Subjects seem to be using the same processes (see high intercorrelations in Table 4.7), why should one type of code provide better access than another? What almost certainly is the answer is in the nature of the organization of the stored materials.

From Table 4.8 there is some initial support for the distance approach, in the final column where it can be seen that though Subjects produce less on average in acoustic, they collectively produce a great deal more of the possible words. This would be in keeping with a theory which proposed that the code information acted as a selector cue for an area in memory and that then words were produced which were close to this area and exhibited the features required (a test which would be applied to the retrieved word). Words arranged closely would be produced together and in predictable sequences, based on their interrelation. Once particular locations based on the initial cue have been seemingly exhausted the search would be switched to a more fruitful area. This type of "random" switching would produce better results in certain situations and would reflect more clearly variables such as set size. In this particular situation, it would mean that whereas more words could be produced in a specific location with a semantic cue, since this is the basis of organization in memory, with the acoustic cue, words are produced by shifting the location and that as this shifting varies from individual to individual so the actual words produced vary considerably. This is also true for visual codes though the real difficulty of the task makes the raw figures unable to illustrate this. A more illustrative way of representing this is to equalise the number of words produced in each of the conditions and then to calculate a figure for the proportion of words produced which were different ones - it can then be seen that the order is visual having the greatest number of different words, followed by acoustic and then semantic.

The clearest way to visualise this is to imagine a "distance" concept, though it is by no means certain that such a system can always be upheld - see the comments concerning Pribram's ideas in the introductory section. In this instance it makes it easier to discuss when a tangible variable such as distance is postulated. Thus the crucial variable in word production is the relative distance of the words being produced - words close together will be produced quickly and easily given that the correct cue for search in that specific area is provided.

The other significant variable to add to the model is "familiarity" which can be conceptualised as a subjective index of word frequency or word usage. The effect of increased familiarity is to reduce the threshold for production of that item. This is in line with a great deal of the findings of researchers in Word Frequency Effects (Broadbent 1972 , Underwood and Schulz, 1960). As a result of these postulates the probability of output of a specific word given search from a particular start point, is directly related to the distance of the word (including paths through mediating items - compare this notion with Kiss's conductance) and to the familiarity of the word.

The other system which is of extreme importance in all this is the Search Mechanism. The formulation of this fundamentally affects the working of the Memory Base part of the model. In this case it is suggested that the search is an all-or-none event, emanating from a point reached by the use of a code or cue and spreading as a "wave" as suggested by Meyer (1973). This does make sense in the light of physiological knowledge, in that the transmission of impulses in the nervous system is an all-or-none affair. Increased search is represented as an increase in the frequency of a search impulse, (i.e. increase in the frequency of initiation of search impulses from the particular starting point) thereby increasing the probability of output of particular words. Every search need not end in success i.e. the output of an appropriate word, as this must depend jointly on the two variables mentioned above, in the Memory Base. However, each succeeding search either by shifting the location of search closer to the word, or by increasing the familiarity value by examination of the word, will increase the probability of output of that word.

This would seem to be quite reasonable and would characterise the situation in a word production task, where more and more words are produced gradually getting more and more unfamiliar. This would not be explained by objective pre-test values of word frequency, since if the word was above threshold it would have an equal probability of output at the beginning of the sequence as at the end of the sequence. This would not appear to be the case - Mednick (1962).

However, an important qualification must be made to this picture of increasing probability of output, and it takes the form of an idea "stolen" from Physiology. This concept is Inhibition. In this case it refers to the decreasing capability for access to available words as a function of the recency of the previous access. Thus, the more the search is initiated in a certain area the greater the build-up of inhibition, leading to words being unavailable. In the experiments reported this can be seen in the fact that though certain subjects produced a large number of responses, there were notable, very common omissions. This is most clearly seen in the category of semantic information, where it is hypothesised that the appropriate words are arranged closely together. For example, subjects, after having given bag, box, trunk, holdall etc., failed to produce "case". There are a great many examples of this in individual subject's productions, where extremely common words, rated as AA in the Thorndike-Lorge count, are not produced, while much less common words appear. It is not clear how the mechanism of inhibition works, but it is fairly certain that something of the sort must exist, if only to inhibit the words which have already been produced and limit the search to words which have not yet appeared.

The nature of the frequency of the search process, must be ascribed to a higher level process, which essentially assesses the value of the task to the subject. This may be in terms of the rewards involved, the punishment inherent in failure, or, as in most Psychological experiments the interest of the subject, as it is clear that there are no differential rewards for better performance. Aspects of this are motivation and attention, and also very importantly the feedback from success or failure in the task. It is suggested that as there is more failure there becomes less pressure from this higher process, less interest and thus in this instance, less search. Theoretically, this takes the form that search decreases as a function of the number of consecutive searches without success. This would mean that towards the end of such a task, as items become more and more difficult to produce, an asymptote is reached, far below the actual capability of the subject in terms of the availability of the words in the set. (This would confirm the views of Tulving and Pearlstone, 1966, and others concerning the notions of availability and accessibility). One indication of this postulate can be examined in the individual latency graphs for word production seen in Appendix 18. It can be seen that items are output in groups of 3, 4 and 5 in close temporal contiguity. What is noticeable about this is that though the gaps between the groups increase as time goes on in the task, the gaps within the groups remain relatively stable. This may well reflect the grouping of items in Memory (in groups of 3 and 4) but it may also signify a decrease in search over time as less and less items are found. This aspect of the model is of a speculative nature and it is appreciated that at present there is no test of the idea, however it does tie the model into other cognitive processes and indicates an area which may have important ramifications for the earlier part of the study.

In summary, the model divides the working of memory into two sections - the Memory Base with the relevant variables of distance and familiarity, and the Search mechanism, an all-or-none phenomenon ruled by a central process reflecting the subject's involvement in the task. Given search from a particular point in memory, output is determined by the variables in the memory base.

It is now appropriate to examine the findings of the final experiments and discuss their precise relation to the model proposed.

The findings of Experiment 4 are in line with the derivations of the model and are explained by the variables in the memory base. However it did raise some questions concerning the relation of the code information and the storage system. These questions were tackled in Experiments 5 and 6.

Experiment 5 examined the time taken to name a word given the code information. The purpose was to compare the various fits of current models relating to the storage system with the model proposed. It was surmised that though a combination of Category Size and Word Frequency could explain the differences in naming time for the various words presented in coded groups, it could not account for the production of these words given the code information. The extremely strong effect of code information throughout this experiment (words are named significantly more quickly when a semantic cue is given than when a physical cue is given, visual or acoustic) can be interpreted with the assumptions about the organization of the memory base, which suggest that the structure of memory is determined by the usefulness of the relation of the words. Thus, as the vocabulary develops words which are used in relation will be stored close together, and words with lesser links will be more distant. It follows from this assumption that words related in meaning will be arranged close to one another - correspondingly, the predictions of the model are that given the correct cue information to begin search at the appropriate point words will be produced or identified as a function of the distance and familiarity inherent in their activity levels. Essentially this would suggest that semantic would be faster than acoustic which would be faster than visual.

One point is clear from this treatment of the code information and this is that the value of the code is to point the search mechanism in the right direction. Its purpose in normal circumstances would be to analyse the incoming information and initiate verification or some form of search in the appropriate area, but in the simple situation presented here it serves merely as a directional arrow (it could be added that in such a test it forms an inefficient director, as subjects without the cue tended to produce faster responses). The point here is that the "code" is nothing mystical, it may simply signify the level of analysis of incoming information, following the normal sensory pattern and leading to an interpretation in the base of information already available (memory). It is easy to see the danger of over emphasising the role of coding as some more recent works have done (Herriot, 1974), for without clearly stating the assumptions underlying the structure of the system which uses these codes to classify and retrieve information, it can become of a similar status to "mental energy" and will "explain" most aspects of memory.

Reasoning as above, the model can account for the fact that retrieval of a word is faster when the subject is given semantic information about the class of words from which it comes, than when he is given acoustic or visual information. Using the concept of familiarity, it also accounts for the difference found between words which had previously been generated by the subject in a similar context and words which had not (Experiment 5). The implication of this finding is that each subject's idiosyncratic organization of words based on his experience should be taken into account in any experiment where there is entry to his memory base.

To support the theory that the code does nothing more than act as a sign-post to the memory system, one test could be made and this was the purpose of the final experiment. It was reasoned that if the code only pointed the way, and the effect manifested was due to the differential arrangements in the Memory Base, then a task which did not involve entry into the Memory Base, but which involved the use of the codes, would not indicate any effect of the different code instructions. This was achieved by asking the subjects to make the decision whether the presented item corresponded to the provided code information. The word need not be identified. The result was that the code effect disappeared, in keeping with the model's predictions.

In some respects the results differ from those which have been previously found e.g. Meyer (1970) showed that there was a category size effect when subjects had to make a category membership decision, which is a task not unlike the one in Experiment 6. However there were one or two differences - Firstly the measure of category size used by Meyer was one of super-ordinates (animals, mammals, living things) which not only increase in actual members but may also require more complex exclusion rules, which thus increase decision time. Experiment 6 had a more basic measure of category size which reflected the number of possibilities available to each individual subject. Secondly, it is possible that the nature of Meyer's task suggested to the subject that he must test the presented information against category exemplars, in which case it would involve entry into the memory base - thus there might be some effect of category size.

The value of Experiment 6 is then in showing that the effects of code found in Experiment 5 can be attributed to the Memory Base, and it also confirms the picture of Codes as directional arrows, used as a starting point in search in the memory system. Thus the model proposed seems to be the most satisfactory explanation of the results found.

A question to be asked of the model is how and to what extent it fits into the pattern of results which have led to the proposals of the various models current in information processing which have been mentioned in the Introduction. The difficulty of this is that most models are based on the "presentation then recall" type of investigation and deal primarily with STM and LTM rather than Very Long Term Memory. However certain derivations can be made from the model proposed here which have relevance to the other systems.

The similarities between the model proposed and the classical systems of Sperling (1962), Waugh and Norman (1967) etc., are on the very basic level of information flow, that the input is processed in a manner dictated by the task requirements and by the subject's previous experience. The differences are in line with present Coding Theory, in the disenchantment with the structures of STM and LTM. Thus there is less emphasis on the "control boxes" seen in the classical system. Concepts such as search are seen as ongoing processes of excitation leading to the raising and lowering of the thresholds of availability, rather than a specific "mechanism" which initiates search. This is in keeping with the Coding Approach, though at the same time there must be certain control processes, which constitute conscious decisions to search for particular features or items but the effect of these is not as pronounced as in the earlier systems. STM thus, becomes a stage in the information flow process, rather than necessarily a storage system - it may correspond only to the completion of certain analysis, which on requesting output produce the characteristic confusions and problems found (Conrad, 1964). That this may represent a level of analysis based on the requirements of the task, can be argued on the basis of Baddeley's (1967) findings that STM does not only involve one type of code, (there may also be an effect of semantic codes, as well as the more normal acoustic). The level of analysis applied will vary according to the task and with the subject involved. This can be clearly seen with deaf children (Conrad, 1972). The model then, attempts to simplify systems such as Norman and Rumelhart (1972) into more basic elements, with a conscious "decision-maker" working on these simple processes. The emphasis is on processes rather than mechanisms.

Other comparisons are relevant. The notions of Episodic and Semantic Memory suggested by Tulving (1972) are not differentiated in the model. What is important to the model is not Episodic and Semantic Memory but rather Episodic and Semantic coding. All memory is episodic in that it involves attributes of the situation of presentation and all the aspects of importance to the subject, which includes meaning as the most important. In certain tasks it may be useful to distinguish the different types of codes which predominate but any implication of differing memory systems is ignored by the model.

Suggestions by Craik and Lockhart (1972) concerning different levels of coding semantic being "deeper" than physical, are quite in keeping with the model at the encoding stage, but are probably not different at the storage stage, being represented as attributes with familiarity and distance parameters. There is some similarity here with the treatment of the episodic-semantic distinction.

Associations do exist in the model and indeed the representation of the distance parameter almost implies an association network. However this is seen only as a convenient form of conceptualisation of the memory base; a system which at present is sufficiently accurate to deal with the results obtained. The notion of association would require the words to be stored as units at fixed distances determined by associative learning. As has been pointed out this may not be a correct picture of the memory base and it is more likely that the structure is akin to that proposed by Pribram (1973). This would mean that items are represented by their attributes (compare this with the formulation derived by Herriot (1974), from the standpoint of coding theory) and the excitation of these attributes contribute to the "wave-form" which when decoded, is the item. The concept of distance becomes a much more complex one, as it is necessary to talk about the closeness to one another of the different attributes. Thus two words are close in memory in the degree to which they have common features (physical, semantic, syntactic etc.) and distant to the extent that their attributes form different wave patterns. This is not as all-inclusive as it may sound as there are indications that certain features are more important in certain task situations - hence the varying emphases on codes for different tasks. All this makes the simple association model inaccurate, and in some circumstances inappropriate, and it is for this reason that the model ultimately does not involve all the assumptions of association.

One system which on the surface at least, is similar, is that of Collins and Quillian (1969) where storage takes the form of a logical network of exemplars, properties and super ordinates. Retrieval or verification is dependent on distance but this distance consists of relationships or properties. Words appear together to the extent that they have similar properties. The system is confined to semantic attributes and to that extent is different from the model proposed. The two systems do attempt to do similar things but problems have arisen with the Collins and Quillian model. Verification of statements concerning words in the system was the task used to test the model and it was found not to be useful for negative statements or statements which were to be rejected as false. Also a basic principle - cognitive economy, where attributes referring to a total class are stored with the superordinate - was found not to hold (Conrad, 1972).

The final question to be asked is where the model stands in relation to coding theory. In a recent statement of the coding approach Herriot (1974) follows the line of Posner and his co-workers, in that active methods of classifying words are used by the subject when the items are presented. These methods take the form of extraction of features, mostly on the lines of physical, name and semantic codes. However, Herriot sees the stored representation as attributes which form a picture of the item - this is similar to the system outlined in previous chapters but his view of word production is somewhat different. He proposes that word production is construction of that item, from the basic given attributes; thus the concept of search is not relevant as the stored representations are attributes rather than the whole words, thus, given certain properties of the word, examples are constructed and presumably tested. The problem here is that the storage of attributes is open to the same disadvantages that are ascribed to the storage of word units. That is, perhaps Herriot's system does not go far enough.

"It is argued that it is mistaken to speak in terms of a lexicon or dictionary of words in the head. Instead, words should be seen as the behavioural products of construction by means of attributes. The construction of words as opposed to non-words has then to be explained in terms of previous learning of certain constructions; that is, of certain ways of combining elements." page 144.

Unless it is emphasised that this system of attributes is completely subjective and based on the subject's experience and his learning of patterns of attributes specific to his learning experience, it is not easy to see how this proposal differs, except in increasing complexity, from the more traditional storage of word units. To deny the search process, is more a denial of a search <u>mechanism</u> since it is clear that even given certain attributes the subject must <u>find</u> others in order to produce a series or group of attributes appropriate for response production. The impression is that like many information processing approaches this one does not go deeply enough into the basis of storage production. Nevertheless it is clearly an approach which goes some way to bridging the gap between classical theory and the model proposed here. There are many similarities especially the notion of attributes acting as the code on input and useful for finding requested items, but it does not go far enough in the multiplication of the storage system, whereby what is stored is the subjective representation of a pattern of attributes and correlates of the item in the presented situation.

Overall, the main difference from previous models of the memory base or Long Term Store, is in the fact that more than semantic coding can be used and it is proposed further that the individual in building up his Long Term storage does so with the codes or attributes which are seen as most useful by him at the time. Thus, there is only partial agreement with Herriot's statement (page 169):

"In LTM tasks, subjects can usually code as much as they wish. They will normally employ semantic codes, since these are better for retrieval purposes." The point being that the subject uses semantic codes not simply because they provide better retrieval cues (this is not a useful explanation, since there could be a large number of ways in which better retrieval could be effected) but because this is the way the memory system has been built, in order to reflect the necessary communication aspects of language. If the individual is a social "animal" his memory system will develop in a similar fashion to others and will rely on semantic appreciation as a recording system best suited to his needs. If for any reason the individual does not have access to this communication function of language, or chooses another, he may develop a completely unique and for normal purposes a completely inappropriate system of access and storage. The two examples of this are: deaf children who have no access to the function of language in the early years and correspondingly have very little development of semantic codes for use in language; an unpublished study of poets by G. Paul indicated that their associations to words and phrases produced significantly more opposite meaning associates than a control group, implying a different relational structure in memory.

Other recent work which has come into the area of Distance models has not gone far enough in exploring the individual parameters involved. Henley (1969) and Rips, Shoben and Smith (1973) illustrate this. Henley (1969) correlates subjects' ratings of dissimilarity of words with a "distance" measure (mean position in the production sequence) and finds a very low non-significant correlation - a result to a large extent, of the blurring of the individual differences which would have readily been available, and also the neglecting of the subjects' use of "familiarity" as a relevant variable, as well as the distance function, in their estimation of similarity. Despite the low correlations, she concludes that there is a common semantic structure available to subjects, a conclusion which would not appear to be realistic in the light of her data.

Rips, Shoben and Smith (1973) support this and maintain that semantic distance effects cannot be readily explained by models in which memory structure mirrors logical structure, a point made here. However, they go on to test their ideas using a verification of logical relations methodology. They do acknowledge that some attribute models such as Meyer's (1970), modified to suggest which attributes are of importance, may account for their data, but are still bound by the notion that it is solely semantic structure and semantic attributes which exist.

The implication is clear: "semantic" memory is not wholly semantic if it is examined in terms of distance relations, and neither is it logical in structure. The proposal is that a concept such as "familiarity" would help to explain the intervening problem of relations in such a distance model.

The outcome of all this is not that the model proposed is the most suitable or accurate one for all circumstances (essentially it has been constructed to deal with the data presented here in a manner illustrative of the underlying principles of this research) but rather that a study of the development of the memory system, given that codes or attributes provide a better fit to the results, should be based on how the individual organises the most efficient use of the system available. It is clear that to a large extent laboratory studies have produced conflicting and confusing implications and it is held here that until a good deal has been explored of the subjective system under-pinning the overt encoding-retrieval system, this confusion will multiply.

IMPLICATIONS FOR EARLIER EXPERIMENTS

The application of the model to the findings of the earlier experiments is not quite as simple as it is for the later experiments, as the type of question the model answers is not the same as the questions that were asked at the outset of the research. Nevertheless there are some interesting aspects of the findings which are of relevance.

The research developed from a simplistic look at original children and their recall of unusual items towards ever-increasing questions of the underlying processes of the storage of words. Experiment 1 attempted to discover general differences in recall for unusual and common words in a group of children who were placed on scales of originality and unusualness of associates. The results were insignificant - there was no broad difference in capacity to recall words which had been generated by the subject to fit the category of common or uncommon words, no difference which could be attributed to differences in originality. At the same time it was discovered that a measure of the subject's performance in generating original words when asked for uncommon words was related to the measures of creativity and also to the unusualness of word associations. This measure then formed the basis of further investigation. The picture that emerged was of the original person as an individual with the same size of "word space" and similar capacities to produce the common words but rather greater access to unusual responses.

Experiment 2 attempted to differentiate between two general models which could describe word production - the evidence tended to support the coding approach rather than a system based on an association network. Also a fairly natural relation between vocabulary and the unusualness of words produced was found. However the most important finding was that though original people could produce more unusual ideas, other people could evaluate them equally well, suggesting that the differences lay on the side of the finding of the appropriate word and not in the evaluation or consequent output.

Experiment 3 confirmed the findings of the first two experiments while clarifying the nature of the originality measure. At the same time it was found that subjects' ability to produce unusual words had improved indicating a better strategy for production. This firmly placed the locus of the originality in the area of encoding or search. Experiment 4 suggested that there were no differences in production according to codes which corresponded to the differences in originality, and this is where the study became more detailed in the study of the memory base.

How can the model add to or explain these findings? As the model stands the most immediate general solution would be to state that as individuals develop their own characteristic word structure, then perhaps those who are creative have simply developed in a different way. This unfortunately cannot be upheld as it was found that original people produced the same responses as other people when asked for common words. At the same time other people can appreciate and evaluate unusualness as accurately as the original people. Both these findings tend to mitigate against great differences of structure in the memory base.

The model has essentially 4 stages (all of which are in operation virtually all the time) and these are Encoding, Search, Evaluation or Testing and Output. Experiment 4 tends to discount a relationship between different codes and originality, thus eliminating the stage of encoding. The previous experiments have discounted the Evaluation section. This leaves only the Search process as the instigator of the divergencies in word production which are seen in output. Support for this idea can be seen in the finding that the original groups were more fluent in their production of words based on code information (Table 4.7), and generally this is a common finding in research on Creativity that creative persons also produce more. The motivation of the creative person has also been found to be higher (Torrance 1963 and others) and the tempting jump to make is to the effect that this higher motivation will result in increased search. The original person will thus initiate more search in the time available than the other person.

The type of motivation referred to here is what would probably be called achievement motivation i.e. it is an indication of the person's aspirations with regard to specific tasks. In a certain sense then, there may be nothing revolutionary in the above contention. Munn (1966) describes the sense of it:-

"... anything that initiates activity, whether external or internal, is motivating. In psychology, however, the terms motivation and motive refer to activation from within the organism. Thus motivated behaviour is internally activated, or at least modified by, internal conditions" (page 151).

It may only be then that what is being said in this instance is that the initiated activity takes the form of a search. However a great deal more is implied - Motivation is a trait which has been identified (Cattell, 1965) and is associated with a great many other aspects of behaviour. It would fit the findings into a general framework of psychological knowledge of behaviour traits of the individual. It would thus seem to satisfy the conditions of more than one level of explanation, and links the predictions of a model of the memory base, derived from experimental evidence and guided by certain findings of physiological psychologists, to the level of abilities and traits. The importance in creativity of the motivational aspects can be seen in the descriptions provided by Roe (1962) and MacKinnon (1962).

One note of qualification must inevitably be placed in this reasoning, and that should be to the effect that what was set out in the proposal of the research, was a study of Originality, namely a creativity defined by the frequency of responses in certain categories. What was eventually studied in depth was a specific small area of word production, for which one can only claim application to Originality in the extraction of verbal material, and it may be acknowledged that there may be a great many areas which are influenced by Originality (Guilford's, 1967, model of the intellect indicates this). It is with this in mind that the issues are clarified.

According to the findings which have been presented here, there are no consistent differences between the person who produces unusual responses in a word production task, in the areas of the use of codes, in retrieval or recall, in evaluation of the uniqueness of items presented. In the model presented there remains only one process which can account for the difference in output, and this is Search. The suggestion of how this works in favour of the more Original person is in terms of increased number of searches for the cued items in the time allowed for the task. This can be identified neatly with Motivation, a factor which has been found to play a role in the make-up of those who are generally better at open-ended tasks. Interpreting this, the proposal is that the crucial difference in Originality (Creativity) lies in Motivation (given the specialist knowledge required in the particular area of application) on the level of overt testable traits and in search on the experimental level and possibly in the excitation of patterns representing items or words, on a more basic level. Appendix 18 gives a more rigorous treatment of the nature of the search process and its temporal parameter, but it can be summarised as being a process whereby success produces further search until the area of investigation is exhausted, and failure reduces the rate of search.

DIRECTIONS FOR THE FUTURE

The two lines of approach relate to the two fields of study - the cognitive process of originality and the study of memory.

(1) Dealing with the former first, it can be seen that the main stumbling block arises because of the fact that only the field of verbal originality is well developed. Even in saying that the results which have been presented have implications for the world of the creative writer or poet, there must be a qualification. Psychologists have as yet, no suitable model of language which could possibly describe the way words are put together to form sentences or how these sentences are processed by the receiver. How can sentences produced by a subject, or recalled by a subject be classified as unusual or original? This highlights the problem in a relatively well studied area, so how can the problem be tackled in the field of music or art? Under these circumstances, experimental research will be limited to the simple type of open-ended tasks presented here or will concentrate on the gross psychometric type of study which may say little about the individual. The development of a suitable system of characterisation of a creative process or simply the production of original works, ideas, stories, even sentences would appear to be a discouraging, long way off. The contention in this thesis is that findings on the basic level as presented here (i.e. originality and fluency of word production) can be

fitted into the framework of cognitive studies as the appropriate models emerge, and that the techniques of investigation as pursued here, with the emphasis on measures such as vocabulary as indicators of the <u>individual's</u> starting point, may provide a mode of study which can be adapted to the changing needs of the experimental situation.

(2) Further studies in the experimental Psychology of memory are suggested in the direction of the individual and the nature of the task involved in the storage and retrieval of items from the vast expanse and resources of memory. This means preferring the studies of production of items to the examination of limited sets of presented items which the subject has to recall or recognise at some accepted interval indicative of short term or long term memory. As also stated in the text above, the establishment of basic measures of the individual's capacity in terms of the words available, the method of organisation which may be idiosyncratic, could be crucial variables in the experimental study of verbal memory. If the development of this type of approach is deemed to have been very slow much of the blame for this may be laid at the door of computer modelling. These models are often based on insufficient knowledge of the base of memory. This occurs simply because in order to explain the findings of presentation-recall studies, the computer program must be given a memory base which can deal with incoming material until it is required and can then be recalled. The problem

here is that this can be done very easily by a computer - after all its recording and storage facilities are the basis of its design and usefulness. Simple representations of stored items can then be made which will allow testing of the factors influencing the presentation-recall paradigm. No further information of the workings of the memory base have been required, and the terms storage and access come to be used only in the computer sense. The basic difficulty in all this is that the human system is not simply a storage system, but is a learning system, a system which reflects experience probably in episodes, in a dynamic, changing way. An illustrative difference may be that a computer will more naturally use a reductive coding system, as in a library, and a human system is much more prone to use elaborative coding. The differences are immense and the limitations imposed by adoption only of the storage attributes of a computer, are pronounced. More information about the human memory base is required.

At the same time all information should be used in the modelling. The suggestion is in line with the statement of Kiss, that it is not necessary to be physiological in dealing with memory, but physiological findings should not be ignored.

Bringing these ideas together the future of research may be towards the examination in detail of word production tasks in those who are recognised as creative, and establishment of the proposed relation with motivation. The key technique in experimental study must be production.

To some extent there are some signs that similar techniques are coming to the fore. Tulving's (1962) notion of subjective organization instigated research of the nature examining the subject's imposed organization, but only in the accepted framework of a presented limited set of items. It might be more realistic or natural to obtain production for the desired attribute and then to work backwards as here, towards presentation and all the relevant measurable variables such as response speed, recognition etc.

In conclusion then, the model developed can account for the results obtained in the reported experiments. It has implications for techniques of study in the field of memory and makes suggestions as to applications to the level of cognitive functioning, via the construct of motivation. Conclusions at present, must be limited to the field of verbal memory representation. How and whether they can be true of other aspects of the memory process, remains to be seen.

Bibliography

ALLEN, M.M. Cueing and retrieval in free recall Journal of Experimental Psychology, 1969, 81, 29-35.

ANDERSON, J.R. FRAN: A simulation model of free recall chapter in G.H. Bower: Psychology of Learning and Motivation, vol. 5, New York: Academic Press, 1972.

ANDERSON, J.R. and BOWER, G.H. Human Associative Memory. Washington D.C., V.H. Winston and Sons, 1973.

ATKINSON, R.C. and SHIFFRIN, R.M. Human Memory: a proposed system and its control processes, in K.W. Spence and J.T. Spence (eds.) The Psychology of Learning and Motivation, vol. 2, New York: Academic Press, 1968.

ATTNEAVE, F. Some informational aspects of visual perception, Psychological Review, 1954, 61, 183-193.

ATWELL, C.R. and WELLS, F.L. Wide Range Vocabulary Test, Psychological Corporation, New York, 1945.

BADDELEY, A.D. Short term memory for word sequences as a function of acoustic, semantic and formal similarity, Quarterly Journal of Experimental Psychology, 1966, 18, 362-66.

BADDELEY, A.D. and DALE, H.C. The effect of semantic similarity on retroactive interference in long- and short-term memory. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 417-20.

BARRON, F. Creativity and Psychological Health. Canada, Van Nostrand, 1963.

BOUSFIELD, W.A. and BARCLAY, W.D. The relationship between order and frequency of occurrence of restricted associated responses. Journal of Experimental Psychology, 1950, 40, 643-47.

BOUSFIELD, W.A. The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 1953, 49, 229-40.

BOUSFIELD, W.A. and COHEN, B.H. The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. Journal of General Psychology, 1955, 52, 83-95.

BOUSFIELD, W.A., COHEN, B.H. and WHITMARSH, G.A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychological Reports, 1958, 4, 39-44.

BOWER, G.H. Multicomponent theory of the memory trace in K.W. Spence and J.T. Spence (eds.) The Psychology of Learning and Motivation, New York: Academic Press, 1967, 229-325.

BOWER, G.H. Stimulus sampling theory of encoding variability. In A.W. Melton and E. Martin (eds.) Coding Processes in Human Memory, Washington, Winston, 1972.

BOWER, G.H., CLARK, M., LESGOLD, A.M. and WINZENZ, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behaviour, 1969, 8, 323-343.

BROADBENT, D.E. Perception and Communication. Pergamon Press, 1958.

BROADBENT, D.E. Decision and Stress. London and New York, Academic Press, 1971.

BROADBENT, D.E. The magic number after 15 years. Conference paper, 1973, now in A. Kennedy and A. Wilkes (eds.) Studies in Long Term Memory, Wiley, 1975.

BROWN, R. and McNEIL, The tip of the tongue phenomenon. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 325-37.

BURT, C. Critical Notice: The psychology of creative ability, British Journal of Educational Psychology, 1962, 32, 292-8.

CATTELL, R.B. The scientific analysis of personality, London, Penguin Books, 1965.

CATTELL, R.B. and DREVDAHL, J.E. A comparison of the personality profile (16 P.F.) of eminent researchers with that of eminent teachers and administrators and of the general population. British Journal of Psychology, 1955, 46, 248-61.

CICIRELLI, V.G. The form of the relationship between creativity, IQ and academic achievement. Journal of Educational Psychology, 1965, 56, 303-8.

CLARK, S.E. Retrieval of colour information from pre-perceptual memory. Journal of Experimental Psychology, 1969, 82, 263-6.

COCHRAN, W.G. and COX, G.M. Experimental Designs, New York, London, Wiley, 1957.

COFER, C.N. Does conceptual organization influence the amount retained in immediate free recall? in B.J. Kleinmuntz (ed.) Concepts and the structure of memory, New York, 1967.

COFER, C.N. and SHEVITZ, R. Word association as a function of word frequency. American Journal of Psychology, 1952, 65, 75-9.

COHEN, B.H. Recall of categorized word lists
Journal of Experimental Psychology, 1963, 66, 227-34.

COHEN, B.H. Some all-or-none characteristics of coding behaviour. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 182-7.

COLLINS, A.M. and QUILLIAN, M.R. Does category size affect categorization time? Journal of Verbal Learning and Verbal Behaviour, 1970, 9, 432-8.

COLLINS, A.M. and QUILLIAN, M.R. How to make a language user, in E. Tulving and W.A. Donaldson (eds.) Organisation and Memory, New York: Academic Press, 1972.

COLTHEART, M. and ALLARD, F. Physical and name codes, paper presented at the meeting of the Psychonomic Society, St. Louis, Nov. 1971.

CONRAD, R. Acoustic confusions in immediate memory, British Journal of Psychology, 1964, 55, 75-84.

CONRAD, R. Short term memory processes in the deaf. British Journal of Psychology, 1970, 61, 179-95.

CONRAD, R. Cognitive economy in semantic memory. Journal of Experimental Psychology, 1972, 92, 149-54.

CORBETT, A., RICE, G. and MILLWARD, R. Category production measures and verification time.
Journal of Verbal Learning and Verbal Behaviour, 1974.

CRAIK, E.J.M. and LOCKHART, R.S. Levels of processing: a framework for memory research.
Journal of Verbal Learning and Verbal Behaviour, 1972, 11, 671-84.

CRONBACH, L.J. Review of J.W. Getzels and P.W. Jackson: Creativity and Intelligence.
American Journal of Sociology, 1962, 68, 278-9.

CROPLEY, A.J. Creativity and Intelligence.
British Journal of Educational Psychology, 1966, 36, 259-66.

CROPLEY, A.J. Creativity. Education Toady Series, London, Longmans, 1969.

CROWDER, R.G. and MORTON, J. Precategorical acoustic storage (PAS).
Perception and Psychophysics, 1969, 5, 365-73.

DE GROOT, A.D. Thought and choice in chess.
The Hague, Mouton, 1965.

DEESE, J. Influence of inter-item associative strength upon immediate free recall.
Psychological Reports, 1959, 5, 305-12.

DONG, T. and KINTSCH, W. Subjective retrieval cues in free recall. Journal of Verbal Learning and Verbal Behaviour, 1968, 7, 813-6.

DORNBUSH, R.L. and WINNICK, W.A. Short term intentional and incidental learning.
Journal of Experimental Psychology, 1967, 73, 608-11.

DOWLING, R.M. and BRAUN, H.W. Retention and Meaningfulness of material.
Journal of Experimental Psychology, 1957, 54, 213-7.

ECCLES, J.C. The behaviour of nerve cells, in G.E.W. Wolstenholme and C.M. O'Connor (eds.) CIBA Foundation symposium, Neurological basis of behaviour, London, Churchill, 1958.

ESTES, W.K. Component and Pattern Models with Markovian Interpretations, in R. Bush and W.K. Estes (eds.) Studies in Mathematical Learning Theory.
Stanford University Press, 1959.

ESTES, W.K. An associative basis for coding and organization in memory, in A.W. Melton and E. Martin (eds.) Coding Processes in Human Memory, Washington D.C. Winston, 1972.

ESTES, W.K. and Da POLITO, F. Independent variation of information storage and retrieval processes in paired-associate learning. Journal of Experimental Psychology, 1967, 75, 18-26.

FATT, P. and KATZ, B. An analysis of the end-plate potential recorded at motor nerve endings Journal of Physiology, 1952, 117, 109-128.

FREUD, S. Leonardo da Vinci and a memory of his childhood, 1910, in Complete Psychological works of Sigmund Freud, J. Strachey (ed.) London, Hogarth Press, 1957.

FREUND, J.S. and UNDERWOOD, B.J. Storage and retrieval cues in free recall learning. Journal of Experimental Psychology, 1969, 81, 49-53.

GETZELS, J.W. and JACKSON, P.W. Creativity and Intelligence, New York, Wiley, 1962.

GHISELIN, B. (ed.) The creative process, University of California Press, 1952.

GLANZER M. and CUNITZ, A.R. Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 351-60.

GLAZE, J.A. The association value of nonsense syllables, Journal of Genetic Psychology, 1928, 35, 255-69.

GOSS, A.E. and NODINE, C.F. Paired Associate Learning: the role of meaningfulness, similarity and familiarization. New York, London: Academic Press, 1965.

GREGG, V. Word frequency, recognition and recall, preprint 1974 for J. Brown, Recognition and Recall. New York, Wiley, 1975.

GUILFORD, J.P. Potentiality for creativity and its measurement in E.F. Gardner, Proceedings of 1962 Invitational Conference on Testing Problems, Princeton N.J., Educational Testing Service, 1963.

GUILFORD, J.P. The nature of human intelligence, New York, McGraw Hill, 1967.

GUILFORD, J.P. and CHRISTENSEN, P.R. A factor analytic study of verbal fluency. Report of the psychology lab, 1956, No.17, Los Angeles, University of Southern California.

GUILFORD, J.P., FRICK, J.W., CHRISTENSEN, P.R. and MERRIFIELD, P.R. A factor analytic study of flexibility in thinking, Report of the psychology lab, 1957, No.18, Los Angeles, University of Southern California.

GUILFORD, J.P. and HOEPFNER, R. Human Intelligence, Academic Press, 1971.

HALL, J.F. Learning as a function of word frequency. American Journal of Psychology, 1954, 67, 138-40.

HARRIS, T.E. The theory of branching processes, Springer-Vetlag, Berlin, 1963.

HENLEY, N.M. A psychological study of the semantics of animal terms, Journal of Verbal Learning and Verbal Behaviour, 1969, 8, 176-84.

HERRIOT, P. Attributes of Memory, London, Methuen, 1974.

HINTZMAN, D.L. Effects of repetition and exposure duration on memory.
Journal of experimental Psychology, 1970, 83, 435-44.

HINTZMAN, D.L. and BLOCK, R.A. Memory Judgements and the effect of Spacing.
Journal of Verbal Learning and Verbal Behaviour, 1970, 9, 561-5.

HOWES, D. On the interpretation of word frequency as a variable affecting speed of recognition.
Journal of Experimental Psychology, 1954, 48, 106-12.

HOWES, D. On the relation between the probability of a word as an association and in general linguistic usage.
Journal of Abnormal and Social Psychology, 1957, 54, 75-85.

HUDSON, L. Contrary Imaginations, London, Methuen, 1966.

HUDSON, R.L. and AUSTIN, J.B. Effect of context and category name on the recall of categorized word lists.
Journal of Experimental Psychology, 1970, 86, 43-7.

HULL, C.L. The meaningfulness of 320 selected nonsense syllables. American Journal of Psychology, 1956, 69, 125-6.

JOHNSON, D.M. and JOHNSON, R.C. and MARK, A.L.
A mathematical analysis of verbal fluency,
Journal of General Psychology, 1951, 44, 121-8.

JOHNSON, S.C. Hierarchical clustering schemes,
Psychometrika, 1967, 32, 241-54.

JOHNSON, N.F. Chunking: Associative chaining versus coding.
Journal of Verbal Learning and Verbal Behaviour, 1969, 8, 725-31.

KAPLAN, I.T., CARVELLAS, T. and METLAY, W. Effect of context on Verbal Recall, J. Verbal Learning and Verbal Behaviour, 1971, 10, 207-12.

KAPLAN, I.T. and CARVELLAS, T. Response probabilities in Verball Recall, Journal of Verbal Learning and Verbal Behaviour, 1969, 8, 344-9.

KINTSCH, W. Recognition and free recall of organised lists. Journal of Experimental Psychology, 1968, 78, 481-7.

KINTSCH, W. Learning, Memory and the Conceptual Processes, New York, Wiley, 1970 (a).

KINTSCH, W. Models for free recall and recognition, in D.A. Norman (ed.) Models of Human Memory, New York, Academic Press, 1970 (b).

KINTSCH, W. and BUSCHKE, H. Homophones and synonyms in short term memory.
Journal of Experimental Psychology, 1969, 80, 403-7.

KISS, G.R. Steps towards a model of word selection, in D. Michie (ed.) Machine Intelligence 4, Edinburgh University Press, 1969.

KISS, G.R. Long Term memory: a state-space approach, British Journal of Psychology, 1972, 63, 327-41.

KISS, G.R. An Associative thesaurus of English, paper presented at conference on current research on aspects of Long Term Memory, University of Dundee, 1973.

KOESTLER, A. The ghost in the machine, London, Hutchinson, 1962.

KOESTLER, A. The act of creation, London, Hutchinson, 1964.

KYLE, J.G. Unpublished M.A. Thesis, Glasgow University, 1970.

LANDAUER, T.K. and FREEDMAN, J.L. Information retrieval from long term memory: category size and recognition time. Journal of Verbal Learning and Verbal Behaviour, 1968, 7, 291-5.

LOEWENTHAL, K. A study of imperfectly acquired vocabulary. British Journal of Psychology, 1971, 62, 225-33.

LOCKHART, R.S. and MURDOCH, B.B. Memory and the theory of signal detection. Psychological Bulletin, 1970, 74, 100-9.

McCORMACK, P.D. Recognition Memory: how complex a retrieval system? Canadian Journal of Psychology, 1972, 26, 19-41.

McGEOCH, J.A. The influence of associative value upon the difficulty of nonsense syllable lists. Journal of Genetic Psychology, 1930, 37, 421-6.

McGEOCH, J.A. Forgetting and the law of disuse. Psychological Review, 1932, 39, 352-70.

MACKINNON, D.W. The nature and nurture of creative talent. American Psychologist, 1962, 17, 484-95.

MAIER, N.R.F., JULIUS, M. and THURBER, J.A. Studies in creativity: individual differences in the storing and utilization of information. American Journal of Psychology, 1967, 80, 492-519.

MAIER, N.R.F. and BURKE, R.J. Studies in Creativity II: influence of motivation in the reorganisation of experience. Psychological Reports, 1968, 23, 351-61.

MAIER, N.R.F., THURBER, J.A. and JULIUS, M. Studies in Creativity III: effect of over learning on recall and usage of information.
Psychological Reports, 1968, 23, 363-8.

MAIER, N.R.F. and THURBER, J.A. Studies in Creativity IV: the appearance of fragmentation and reorganisation in meaningfully associated units.
Psychological Reports, 1968, 23, 979-90.

MAIER, N.R.F., THURBER, J.A. and JANZEN, J.C. Studies in Creativity V: the selection process in recall and in problem solving situations.
Psychological Reports, 1968, 23, 1003-22.

MALTZMAN, I. On the training of originality,
Psychological Review, 1960, 67, 229-42.

MALTZMAN, I., BELLONI, M. and FISHBEIN, M. Experimental studies of associative variables in originality.
Psychological Monographs, 1964, 78, No.3.

MANDLER, G. Organisation and recognition, in E. Tulving and W. Donaldson (eds.) Organisation of memory, New York, Academic Press, 1972.

MANDLER, G. and PEARLSTONE, Z. Free and constrained concept learning and subsequent recall.
Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 126-31.

MANDLER, G., PEARLSTONE, Z. and KOOPMANS, H.S. Effects of organisation and semantic similarity on recall and recognition. Journal of Verbal Learning and Verbal Behaviour, 1969, 8, 410-23.

MATTHEWS, W.A. and MANASSE, K. Associative factors in free recall. Quarterly Journal of Experimental Psychology, 1970, 22, 177-84.

MEDNICK, S.A. The associative basis of the creative process. Psychological Review, 1962, 69, 220-32.

MELTON, A.W. Implications of short term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behaviour, 1963, 2, 1-21.

MENDELSOHN, G.A. and GRISWOLD, B.B. Differential use of incidental stimuli in problem solving as a function of creativity. Journal of Abnormal and Social Psychology, 1964, 68, 431-6.

MEYER, D.E. On the representation and retrieval of stored semantic information. Cognitive Psychology, 1970, 1, 242-99.

MEYER, D.E. Correlated operations in searching stored semantic categories. Journal of Experimental Psychology, 1973, 99, 124-33.

MEYER, D.E. and SCHVANEVELDT, R.W. Facilitation in recognising pairs of words: evidence of a dependance between retrieval operations. Journal of Experimental Psychology, 1971, 90, 227-34.

MEYER, D.E., SCHVANEVELDT, R.W. and RUDDY, M.G. Activation of lexical memory, paper presented at meeting of Psychonomic Society, St Louis, November 1972.

MILLER, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

MILLER, G.A. Psycholinguistic approaches to the study of communication, in D.L. Arun Journeys in Science. Albuquerque, University of New Mexico Press, 1967.

MILLWARD, R. Category Production measures and verification times. Conference paper 1973, in A. Kennedy and A. Wilkes Studies in Long Term Memory, Wiley, 1975.

MILNER, B. Brain mechanisms suggested by studies of temporal lobes, in F. Darley (ed.) Brain mechanisms underlying speech and language, New York, Grane and Stratton, 1967.

MORTON, J. The effects of context on the visual duration threshold for words.
British Journal of Psychology, 1964, 55, 165-80.

MORTON, J. Repeated items and decay in memory, Psychonomic Science, 1968, 10, 219-20.

MORTON, J. Interaction of information in word recognition, Psychological Review, 1969, 76, 165-78.

MORTON, J. A functional model for memory, in D.A. Norman (ed.) Models of Human Memory, New York: Academic Press, 1970.

MORTON, J. and BROADBENT, D.E. Passive versus active recognition or is your homunculus really necessary? in W. Wathan-Dunn (ed.) Models for the perception of speech and visual form. Cambridge, Mass, MIT Press, 1967.

MORTON, J. and HOLLOWAY, C.M. Absence of a cross-modal suffix effect in short term memory.
Quarterly Journal of Experimental Psychology, 1970, 22, 167-76.

MUNN, N.L. Psychology: the fundamentals of human adjustment. Boston, Houghton Mifflin, 1966.

MURDOCK, B.B. The serial position effect of free recall, Journal of Experimental Psychology, 1962, 64, 482-8.

MURDOCK, B.B. Effects of a subsidiary task on short term memory. British Journal of Psychology, 1965, 56, 413-9.

MYDEN, W. Interpretation and evaluation of certain personality characteristics involved in creative production. Perceptual Motor Skills, 1959, 9, 139-58.

NOBLE, C.E. An analysis of meaning, Psychological Review, 1952, 59, 421-30.

NOBLE, C.E. and McNEELY, D.A. The role of meaningfulness (m) in paired-associate verbal learning. Journal of Experimental Psychology, 1957, 53, 16-22.

NOBLE, C.E., STOCKWELL, F.E. and PRYER, M.W. Meaningfulness (m') and association value (a) in paired associate syllable learning. Psychology Report, 1957, 3, 441-52.

NORMAN, D.A. (ed.) Models of Human Memory, New York: Academic Press, 1970.

NORMAN, D.A. and WICKELGREN, W.A. Short term recognition memory for single digits and pairs of digits. Journal of Experimental Psychology, 1965, 70, 479-89.

NORMAN, D.A. and RUMELHART, N. A system for perception and memory, in D.A. Norman (ed.) Models of Human Memory. New York: Academic Press, 1970.

PAIVIO, A. Imagery and Verbal processes, New York: Holt, Rinehart and Winston, 1971.

PAIVIO, A., YUILLE, J.C. and MADIGAN, S.A. Concreteness, imagery and meaningfulness values for 925 nouns. Journal of Experimental Psychology, Monograph Supplement, 1968, 76, No.1 part 2.

PALERMO, D.S. and JENKINS, J.J. Word association norms, Minneapolis, University of Minnesota Press, 1964.

PATRICK, C. Creative thought in Poets, Archives of Psychology, 1935, No.178.

PATRICK, C. Creative thought in Artists, Journal of Psychology, 1937.

PAUL, G. Word associations in poets, Unpublished M.A. Thesis, Glasgow University, 1968.

PETERSEN, L.R. and PETERSEN, M.J. Short term retention of individual verbal items, Journal of Experimental Psychology, 1959, 58, 193-8.

PETTIGREW, T.F. The measurement and correlates of category width as a cognitive variables. Journal of Personality, 1958, 26, 532-44.

POLLACK, I., RUBINSTEIN, H. and DECKER, L. Intelligibility of known and unknown message sets. Journal of the Acoustic Society of America, 1959, 31, 273-

POLLIO, H.R. Composition of associative clusters, Journal of Experimental Psychology, 1964, 67, 199-208.

POLLIO, H.R. Associative structure and verbal behaviour, in T.R. Dixon and D.L. Horton (eds.) Verbal behaviour and General behaviour theory, Englewood Cliffs, N.J. Prentice Hall, 1968.

POSNER, M.I. Abstraction and the process of recognition, in G.H. Bower and J.T. Spence (eds.) The Psychology of Learning and Motivation, vol.3, New York: Academic Press, 1969.

POSNER, M.I. Co-ordination of internal codes, in W.G. Chase (ed.) Visual information processing, New York: Academic Press, 1973.

POSNER, M.I. and BOIES, S.W. Components of Attention, Psychological Review, 1971, 78, 391-408.

POSNER, M.I. and KLEIN, R. On the function of consciousness paper presented to the fourth conference on Attention and Performance. Boulder, Colorado, August 1971.

POSNER, M.I. and WARREN, R.E. Traces, concepts and conscious constructions, in A.W. Melton and E. Martin (eds.) Coding Processes in Human Memory, Washington D.C. Winston, 1972.

POSTMAN, L. Short term memory and incidental learning, in A.W. Melton (ed.) Categories of Human Learning, Academic Press, 1964.

PRIBRAM, K.H. Languages of the Brain, New Jersey, Prentice-Hall, 1971.

RIPS, L.J., SHOBEN, E.J. and SMITH, E.E. Semantic Distance and the verification of semantic relations, Journal of Verbal Learning and Verbal Behaviour, 1973, 12, 1-20.

ROE, A. The psychology of the scientist, Science, 1961, 134, 456-59.

ROSENBLATT, F. A probabilistic model for information storage and organisation in the brain. Psychological Review, 1958, 65, 386-407.

RUSSELL, W.A. and JENKINS, J.J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff association test.
Technical Report No.11, University of Minnesota, 1954.

SCHEFFE, H.A. A method for judging all contrasts in the analysis of variance, Biometrika, 1953, 40, 87-104.

SCHULMAN, A.I. Recognition memory for targets from a scanned word list.
British Journal of Psychology, 1971, 62, 335-46.

SCHULMAN, A.I. Encoding processes and the memorability of events, paper presented at conference on Current Research in Aspects of Long Term Memory, University of Dundee, July 1973.

SHALLICE, T. and Warrington, E.K. Independent functioning of verbal memory stores: a neurophysiological study. Quarterly Journal of Experimental Psychology, 1970, 22, 261-73.

SHANNON, C.E. The mathematical theory of communication, Urbana, University of Illinois Press, 1948.

SHAPIRO, R.J. Creative Research Scientists,
National Institute for Personnel Research, 1968.

SHIFFRIN, R.M. Search and Retrieval Processes in Long Term Memory, Technical Reports No.137, Institute for Mathematical Studies in Social Sciences, Stanford University, 1968.

SHIFFRIN, R.M. Memory Search, in D.A. Norman (ed.)
Models of Human Memory, New York: Academic Press, 1970.

SHIFFRIN, R.M. and ATKINSON, R.C. Storage and retrieval in long term memory. Psychological Review, 1969, 76, 179-93.

SIEGEL, S. Non-parametric Statistics, New York: Wiley, 1956.

SIMPSON, R.M. Creative Imagination, American Journal of Psychology, 1922, 33, 234-43.

SLAMECKA, N.J. An examination of trace storage in free recall. Journal of Experimental Psychology, 1968, 76, 504-513.

SLAMECKA, N.J. Testing for associative storage in multitrial free recall. Journal of Experimental Psychology, 1969, 81, 557-561.

SPEARMAN, C. The abilities of Man, New York: Macmillan, 1927.

SPERLING, G.A. The information available in brief visual presentation, Psychological Monographs, 1960, 74, 11, Whole No. 498.

SPERLING, G.A. A model for visual memory tasks, Human Factors, 1963, 5, 19-30.

SPERLING, G.A. Successive approximations to a model for short term memory. Acta Psychologica, 1967, 27, 285-292.

STEENBERG, S. High Speed scanning in human memory, Science, 1966, 153, 652-654.

STERNBERG, S. Memory scanning: mental processes revealed by reaction time experiments. American Scientist, 1969, 57, 421-457.

TAYLOR, C.W. and BARRON, F. Scientific Creativity, New York: John Wiley, 1963.

THORNDIKE, R.L. The measurement of creativity, Teachers College Record, 1963, 64, 422-424.

THORNDIKE, R.L. and LORGE, Teachers Word book of 30,000 words, Teachers College, New York: University of Columbia.

TORRANCE, E.P. Education and the Creative Potential, Minneapolis: University of Minnesota Press, 1963.

TORRANCE, E.P. Torrance tests of creative thinking: norms-technical manual. Princeton, N.J. Personnel Press, 1966.

TULVING, E. Subjective organization in free recall of unrelated words. Psychological Review, 1962, 69, 344-354.

TULVING, E. Intra trial and intertrial retention: notes toward a theory of free-recall of verbal learning. Psychological Review, 1964, 71, 219-237.

TULVING, E. Subjective organization and effects of repetition in multi-trial free recall learning. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 193-197.

TULVING, E. Episodic and semantic memory, in E. Tulving and W.A. Donaldson (eds.) Organisation of Memory, New York: Academic Press, 1972.

TULVING, E. and PEARLSTONE, Z. Availability versus Accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 381-391.

TULVING, E. and THOMSON, D.M. Retrieval Processes in recognition memory: effects of associative context. Journal of Experimental Psychology, 1971, 87, 116-124.

TURVEY, M.T. and KRAVETZ, Retrieval from iconic memory with shape as the selection criteria, Perception and Psychophysics, 1970, 8, 171-172.

UNDERWOOD, B.J. Verbal learning in the educative processes, Harvard Education Review, 1959, 29, 107-117.

UNDERWOOD, B.J. Attributes of Memory, Psychological Review, 1969, 76, 559-573.

UNDERWOOD, B.J. and SCHULTZ, R.W. Meaningfulness and Verbal Learning, Philadelphia: Lippincott, 1960.

VON WRIGHT, J.M. Selection in visual immediate memory, Quarterly Journal of Experimental Psychology, 1968, 20, 62-68.

WALLACH, M.A. Creativity, in Mussen (ed.) Manual of Child Psychology, vol.1, N.Y.: Wiley, 1970.

WALLACH, M.A. and KOGAN, N. Modes of thinking in young children: a study of the creativity-intelligence distinction. New York: Holt, Rinehart and Winston, 1965.

WALLAS, G. The art of thought, London: Jonathan Cape, 1926.

WARD, W.C. Creativity and impulsivity in kindergarten children. Doctoral Dissentation, Duke University, Durham N.C., 1966.

WAUGH, N.C. and NORMAN, D.A. Primary Memory, Psychological Review, 1965, 72, 89-104.

WICKELGREN, W.A. Time, interference and rate of presentation in short-term recognition memory for items. Journal of Mathematical Psychology, 1970, 7, 219-235.

WILKINS, A.J. Categorization time and category size, Journal of Verbal Learning and Verbal Behaviour, 1971, 10, 383-385.

WILLSHAW, D.J., BUNEMAN, O.P. and LONGUET-HIGGINS, H.C. Non-Holographic Associative Memory, Nature, 1969, 222, 960-962.

WILSON, R.C., GUILFORD, J.P., CHRISTENSEN, P.R. and LEWIS, D.J. A factor analytic study of creative-thinking abilities, Psychometrika, 1954, 19, 297-311.

WOOD, G. Organization, large memory units and free recall. Journal of Verbal Learning and Verbal Behaviour, 1971, 10, 52-56.

YAMAMOTO, K. Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. British Journal of Educational Psychology, 1965, 35, 300-305.

YAMAMOTO, K. and CHIMBIDIS, M.E. Achievement, intelligence and creative thinking in fifth-grade children: a correlational study. Merrill-Palmer Quarterly, 1966, 12, 233-241.

YATES, F. The analysis of replicated experiments when the field results are incomplete. Emp. Journal of Experimental Agriculture, 1933, 1, 129-42.

APPENDICES

Instructions to Subjects

Experiment 1.

Each child was provided with a booklet with the numbers 1 to 30 printed at intervals on the pages, 3 to a page, as indications of the different words to which they were to make associations - thus there was plenty of space for the associations to each word to be written in.

"Will you put your name, class and date of birth on the back. I have come here to ask you to help out with a study. It has to do with the ideas you have about certain things. There are 3 main parts and I'll explain each one as we come to it." (check each has a script and a pencil)

"Well, the first part will be split up into 2 sections between which you will have a short break. What is going to happen is that you will be shown a key word and what you have to do is to write down what the word reminds you of, what it makes you think of. As soon as you have written down your first idea, I want you to look up, look back at the key word and see whether it makes you think of something else, something different. You then write down that idea and continue like that till you run out of ideas."

"The key will be written on a card which I will hold for you to see. You will have 40 seconds for each word. Oh, just one last thing before we start, will you put a), b), c), etc. for each different idea that you put down."

"Is that clear now? Are there any questions? O.K. then let's begin."

After the first 15 words were completed there was a 5 minute break during which time the first part of the booklet, numbers 1 to 15 was collected and the second half 16 to 30, was given out again asking subjects to write their names. Another 15 key words were presented for the same length of time and the booklets were collected.

(2) Each subject was provided with a set of 20 cards 3" x 2" tied with an elastic band.

"Undo the elastic band from the cards and count them. Make sure there are 20 there. Put them in 2 piles of 10. First of all, take one pile and on each card of that pile I want you to write one word. The word must have 2 syllables, that is like "breakfast" or "morning" and it must be a noun. Now the important thing is that the word has to be unusual, it must be one which we don't often come across. Each word must be different. Try to choose different categories. So, one word on each card in the first pile and the word has to be unusual and have 2 syllables and be a noun. After you write a word turn the card over face down. O.K. go ahead." Allow 10 minutes.

"Stop now - in each corner I want you to put the capital letter "A". "A" on "every card".

Now take the second pile. I want you to do the same except this time you must choose words which are common, words which we came across all the time. Again one word on each card and then turn it over. The word must have 2 syllables and be a noun - remember it must be a usual word. O.K. go ahead."

After 10 minutes:

"On each card in this pile, will you put the letter "B" in the corner. Finally on the last card, on the back, will you put your name. Put the cards together and shuffle them, mix them up completely. I'll bring round the next part."

For the final part of this session the subject was provided with a booklet of 20 pages with a random ordering (unique to each subject) of a series of nonsense syllables printed one to a page on the left-hand side.

"Have you shuffled the cards well? Turn them face down on the desk. Open the booklet. On the left-hand side you will see a word, a 3-letter word. What you have to do is to put each one of your words next to one of these words (on each page of the booklet). When I tell you to start I want you to turn the first card over, and write it across from the 3-letter word on the first page of the booklet. Then underneath it, in the booklet, I want you to copy the 3-letter word which goes with it."

(Demonstrate on blackboard)

"You have then to try to think of these 2 words together. Once you have done this place the card at the bottom of the pile, turn over the page and do the same thing for the next card and the next page. Try and think of the 2 words together and then move on to the next one ... and so on till the end. When you come to the end put the elastic band round the cards, shut the booklet and put your name on the back. Leave them on your desk and they will be collected."

In the second session the subject was presented with a blank sheet of paper.

"On the sheet in front of you, put your name, at the top. You will remember that this morning you had to fill in a booklet and write a word which you had thought up above a 3-letter word. All you have to do on this sheet of paper is to remember these words, both your word and the 3-letter word which went with it, and write them down together in a list. Try to remember the ones which went together and write them down. O.K. go ahead."

After a few minutes and especially if the subjects could not pair any of the words and the nonsense syllables.

"If you can't remember the ones which went together, write down the ones "you can recall alone."

Subjects were stopped after 15 minutes and the sheets of paper collected. Booklets consisting of 4 pages were distributed. On the first 2 pages were series of circles 1" in diameter, with the instructions at the top of the page to draw an object which had a circle as the main part.

"Put your name, class and date of birth on the back of the booklet. Don't turn the booklets over yet.

The items here are designed to see if you can use your imagination, to see how well you can think. We want to see how many ideas you have about different things. This is not a test, since there are no right or wrong answers, but there will be a time limit for each question, so that you should use your time well and work as rapidly as you can with comfort.

Put down every idea you have no matter how silly it sounds to you. Don't turn the page until you are told to do so, or unless it says so at the bottom of the page. O.K. turn the booklet over.

See how many objects you can sketch which have a circle as the main part. Just use a few lines on circles to identify your ideas which might start wheel, tyre, steering wheel and so on. Your lines may be inside or outside the circle, or both inside and outside. Write the name of the object you have sketched underneath it.
All right go ahead."
10 minutes are allowed.

"Stop, pencils down. Turn over the page. On this page and the next one are 3 questions: What would happen if man could be invisible at will? What would happen if a hole could be bored through the earth? What would happen if the language of the birds and animals could be understood by man?"

"Now write down all the consequences which you think there would be if these things came about. Write the answers to each one under the question on the page. All right, go ahead."

After 5 minutes

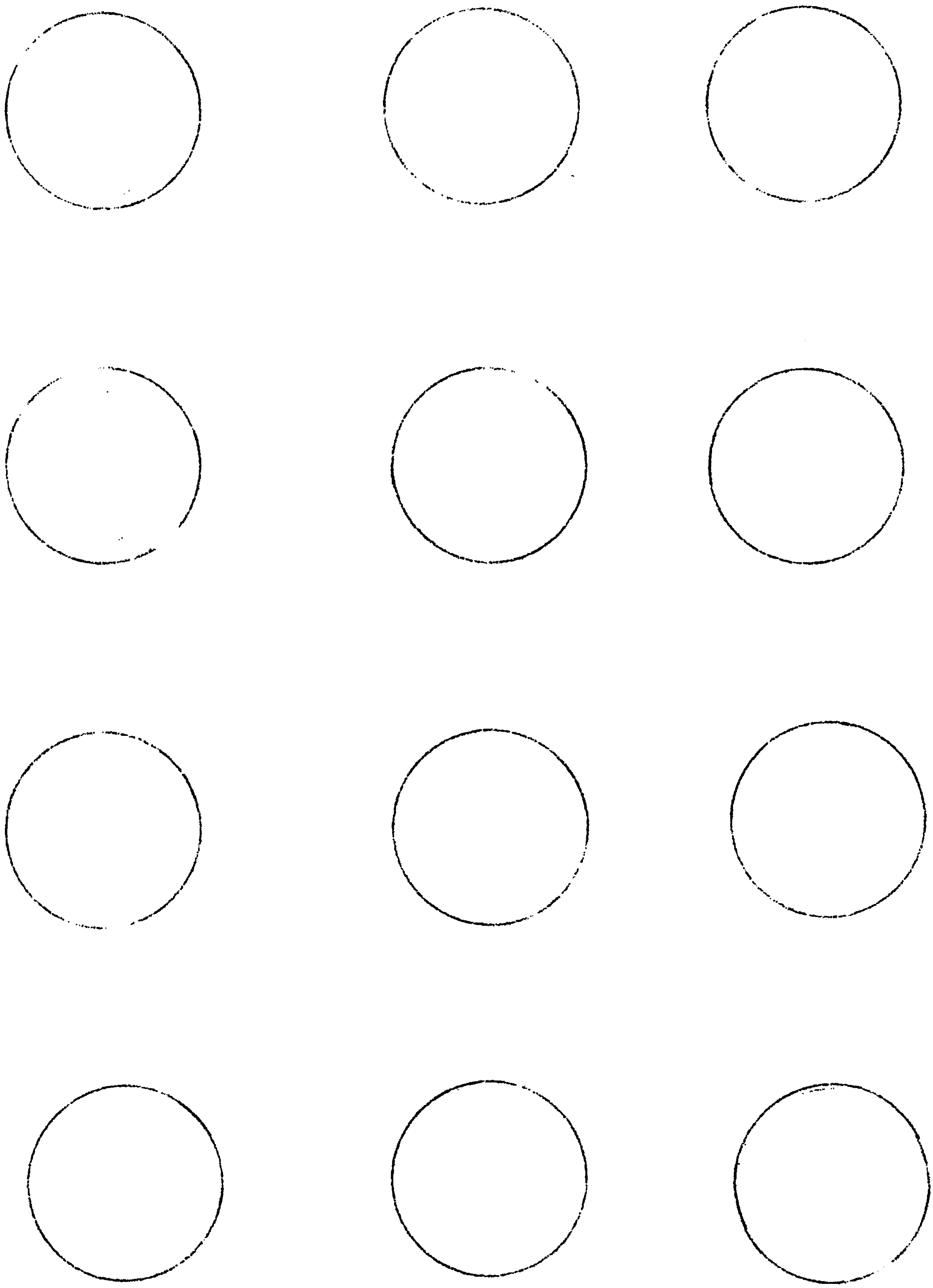
"Stop, pencils down. Make sure you have put name, class and date of birth on the back."

Appendix 1.

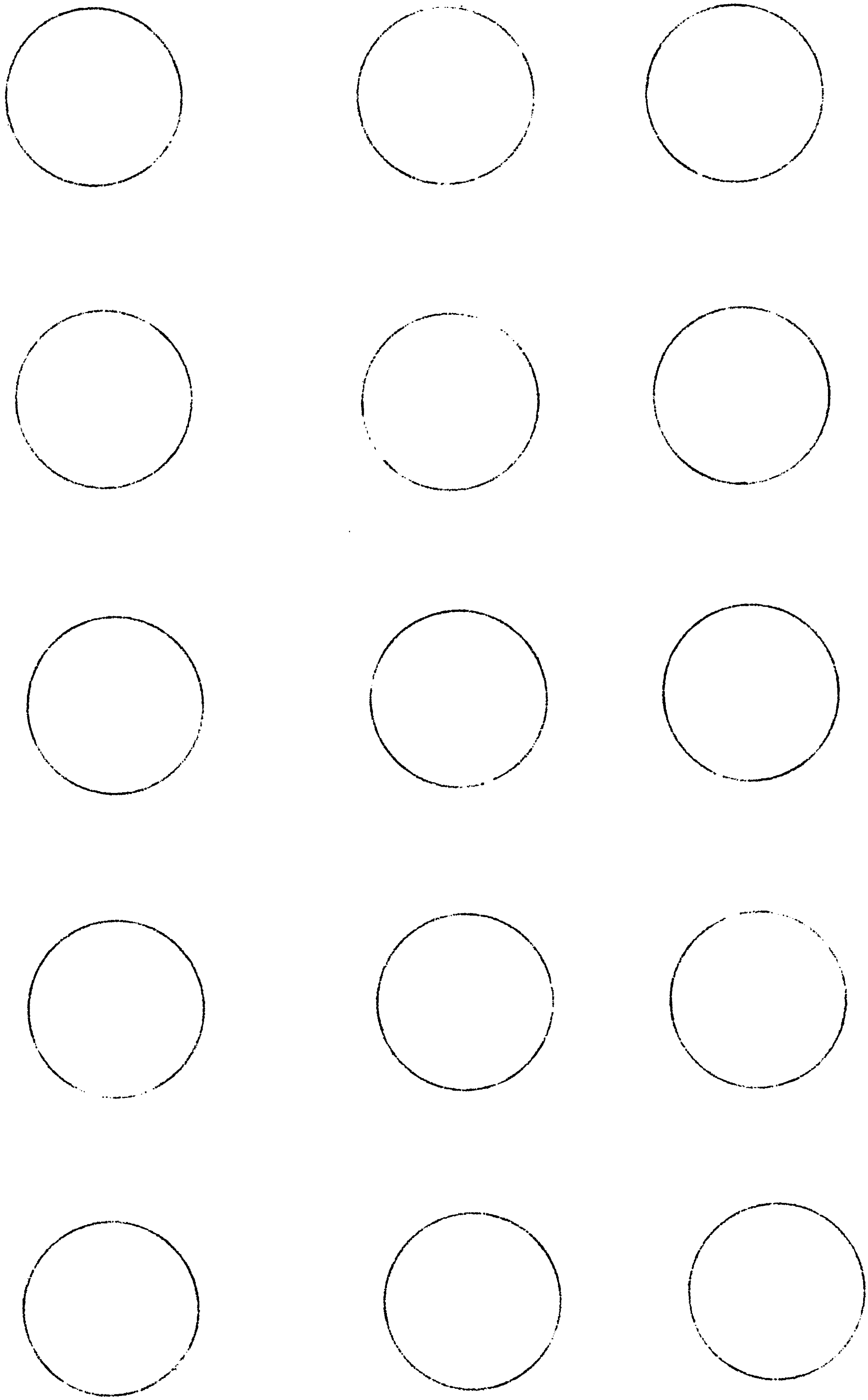
Key Words presented for Continuous Association.

1. Engine
2. Wisdom
3. Forest
4. Oyster
5. Panic
6. Leather
7. Bullet
8. Drama
9. Goodness
10. Region
11. Muscle
12. Author
13. Picnic
14. Circus
15. Tourist
16. Elbow
17. Tennis
18. Platform
19. Absence
20. Budget
21. Acre
22. Device
23. Response
24. Accent
25. Navy
26. Presence
27. Candle
28. Adult
29. Message
30. Oven

Draw as many things as you can which have a circle as the main part.
Write the name of the object underneath.



Turn Over



Do not turn page until told to do so

What would happen if man could be invisible
at will?

What would happen if a hole could be bored
through the earth?

Turn Over

What would happen if the language of birds and animals could be understood by man?

Appendix 2

Raw Data for Subjects in Experiment 1.

Subject No.	% Recall of A words	% Recall of A + nonsense syllable	% Recall of B words	% Recall of B + nonsense syllable	% Recall of nonsense syll.	Unusualness Ratio	Circles	Consequences	Unusualness of A words	Unusualness of B words
1.	60.0	0	75.0	12.5	30	23.13	13.71	44.40	80.0	60.73
2.	77.78	33.33	33.33	0	25	30.45	30.27	33.30	86.1	50.0
3.	66.67	11.11	42.86	14.29	45	11.83	16.65	11.10	58.33	46.43
4.	66.67	22.22	77.78	33.33	50	10.17	6.24	33.30	72.23	44.45
5.	60.0	0	66.67	16.67	30	26.19	21.19	66.6	90.0	41.68
6.	77.78	44.44	87.5	50.0	45	12.01	16.65	11.10	58.33	34.38
7.	100.0	0	50.0	12.5	30	19.85	3.70	33.30	62.5	53.13
8.	33.33	11.11	62.5	25.0	40	27.01	44.4	33.3	77.78	65.63
9.	40.0	0	37.5	37.5	45	29.54	35.86	66.6	90.0	62.5
10.	66.67	22.22	42.86	14.29	25	23.77	22.2	61.84	80.55	61.85
11.	66.67	22.22	28.59	0	40	29.95	16.65	41.62	69.45	60.73
12.	44.44	11.11	44.44	33.33	35	23.52	29.97	33.3	44.45	58.33
13.	75.0	25.0	75.0	25.0	30	13.39	10.25	22.2	56.25	68.75
14.	42.86	0	66.67	22.22	20	43.27	21.05	19.98	78.58	55.55
15.	30.0	10.0	71.43	71.43	45	30.16	18.5	38.85	80.0	60.73
16.	70.0	10.0	42.86	28.57	50	31.98	18.16	42.81	92.5	57.15
17.	20.0	0	50.0	0	0	35.63	10.0	55.5	100.0	62.5
18.	75.0	0	40.0	20.0	25	26.36	13.32	13.32	59.38	62.5
19.	77.78	11.11	75.0	12.5	20	21.92	43.29	38.06	93.68	62.5
20.	42.86	14.29	33.33	11.11	20	30.05	14.58	37.46	85.73	47.23
21.	87.5	50.0	62.5	0	35	26.75	15.77	27.75	65.63	65.63
22.	62.5	0	50.0	0	35	15.55	23.79	44.4	71.88	79.18
23.	44.44	0	83.33	0	0	21.07	22.91	33.3	94.45	41.68
24.	0	0	20.0	0	0	28.42	16.62	24.97	85.0	50.0
25.	57.14	14.29	50.0	12.5	50	38.48	47.87	44.4	100.00	75.0
26.	33.33	0	50.0	12.50	15	24.34	14.27	39.96	62.5	59.38

Appendix 2 Contd.

Raw Data Continued

Subject No.	% Recall of A words	% Recall of A+ nonsense syllable	% Recall of B words	% Recall of B+ nonsense syllable	% Recall of nonsense syll.	Unusualness Ratio	Circles	Consequences	Unusualness of A words	Unusualness of B words
27.	57.14	0	33.33	16.67	40	18.66	23.31	29.00	82.15	29.18
28.	80.0	60.0	66.67	11.11	50	35.38	29.14	66.60	90.0	47.23
29.	42.86	0	55.56	11.11	25	9.16	13.32	49.95	71.43	63.9
30.	57.14	42.86	87.50	50.0	50	16.86	12.11	16.65	78.58	46.88
31.	42.86	28.57	37.50	12.50	45	30.02	7.40	0	67.85	75.0
32.	62.50	0	57.14	0	40	19.61	3.70	55.50	65.63	46.43
33.	55.56	0	33.33	16.67	45	27.75	19.03	41.62	63.9	58.33
34.	66.67	0	40.0	0	20	28.18	19.03	55.50	83.33	42.5
35.	85.71	0	62.50	12.50	35	25.84	16.65	19.03	67.85	53.13
36.	100.0	22.22	22.22	0	35	21.86	12.95	22.20	55.55	61.1
37.	33.33	0	80.0	20.0	45	20.81	16.65	22.20	75.0	40.0
38.	60.0	30.0	71.43	14.29	30	12.65	33.30	26.64	60.0	42.85
39.	42.86	0	44.44	11.11	30	17.65	0	33.30	46.43	63.9
40.	60.0	40.0	100.0	28.57	40	49.43	28.18	41.62	65.0	42.85
41.	16.67	16.67	14.29	14.29	30	21.30	45.19	24.97	87.5	50.0
42.	66.67	44.44	100.0	28.57	65	15.88	23.31	11.10	52.78	28.58
43.	40.0	0	42.86	28.57	15	21.06	12.11	22.20	60.0	57.15
44.	33.33	11.11	28.57	0	65	33.22	26.16	19.03	52.78	70.83
45.	14.29	0	16.67	16.67	10	11.53	27.06	33.30	64.28	70.83
46.	33.33	16.67	85.71	14.29	20	23.56	28.18	19.98	50.0	32.15
47.	57.14	28.57	77.78	11.11	15	14.91	31.66	19.98	85.73	77.78
48.	75.0	0	70.0	10.0	10	24.29	33.30	41.62	87.5	57.5
49.	0	0	0	0	15	14.76	9.99	8.32	43.75	50.0
50.	60.0	20.0	40.0	0	5	20.64	30.94	26.64	50.0	65.0
51.	75.0	0	44.44	11.11	15	16.87	19.42	59.94	75.0	61.1

Appendix 3

"List of words generated as Unusual"

words	frequency of generation	Thorndike-Lorge frequency
actor	1	28
airport	1	3
anthrax	2	1
antique	1	9
armpit	1	1
atom	2	8
austin	1	3
balloon	2	17
ballot	1	14
banquet	1	18
barmaid	1	Y
batman	2	
beauty	1	AA
beaver	1	23
bedroom	1	35
beehive	1	1
belly	1	10
beltane	1	
biscuit	2	14
bismuth	1	X
blackbeard	1	
blackboard	2	9
bladder	1	2
blantyre	1	
breakfast	1	A
briefcase	1	
brothel	1	X
budgie	6	
bureau	2	44
butter	1	AA
buttocks	2	1
cabbage	5	16
cannon	1	21
carol	1	3
carriage	1	46
carrot	1	9
cello	1	X
celtic	3	2
chicken	1	A
china	1	A
chipmunk	1	7
chopper	1	2
clamour	2	14
classroom	1	3
clover	1	12
cobbler	3	8
coffin	1	11
college	1	AA
convoy	1	3
cornflake	1	
coward	1	22

Appendix 3 Contd.

"List of words generated as Unusual"

word	frequency of generation	Thorndike-Lorge frequency
crankshaft	1	
crescent	1	6
croquet	1	1
crumpet	4	Y
cymbal	1	3
dalton		2
damask	2	4
dancing		
dandruff	1	X
dashboard	1	X
demon	1	8
devil	1	A
dinner	1	AA
donkey	2	16
doormat	1	
dragon	2	22
drawbridge	2	3
drifter	1	
egghead	1	
evening	1	AA
exam	2	28
eyeball	3	1
eyebrow	2	12
eyelid	1	9
fairy	2	35
female	1	38
firework	1	3
fisher	1	8
fishman	3	
flower	1	AA
football	4	26
forenoon	1	9
fuel	1	21
garage	1	14
gargoyle	1	X
gatepost	1	X
gecko	1	Y
giraffe	1	1
goblin	1	14
grapeshot	1	
groundhog	2	
guitar	3	4
guru	1	
haddock	1	1
hades	1	2
haggis	6	
haircut	3	
halter	1	5
hammer	1	34
handle	1	A
handwork	1	X

Appendix 3 Contd.

"List of words generated as Unusual"

word	frequency of generation	Thorndike-Lorge frequency
harness	1	29
hatchet	3	8
haystack	1	1
headstone	1	X
highway	1	34
hobo	1	1
hopscotch	1	
hoover	1	35
hornet	6	3
horses	1	A
hotspur	3	X
houseboat	1	X
howkin	1	
hudson	1	30
inkwell	2	X
ishmael	1	1
island	1	AA
islet	1	3
jackboot	1	X
jackdaw	1	1
jacket	1	22
jaundice	1	1
jodhpur	1	
junction	1	8
juno	1	7
kennel	2	6
kestrel	1	
knickers	1	
labour	1	AA
lampost	3	X
lamprey	1	X
lemming	2	
lettuce	1	12
levi	4	1
levite	6	1
libel	1	2
lilo	1	
lion	1	A
loony	1	X
lumber	2	34
mangle	1	5
mango	1	1
marine	3	17
master	1	AA
matter	3	AA
mermaid	1	2
metal	1	A
meter	2	12
migraine	1	
mildew	1	3
mirage	1	1

Appendix 3 Contd.

"List of words generated as Unusual"

word	frequency of generation	Thorndike-Lorge frequency
missile	1	2
mobile	1	2
money	1	AA
monkey	1	23
mongol	1	2
mongoose	1	
moron	1	X
mural	2	3
music	1	AA
musket	1	8
neutron	1	
nigger	1	2
nightime	1	1
nostril	2	15
odour	2	27
onion	1	25
orgasm	1	
outcry	2	7
pakie	1	
panties	1	
parcel	1	23
parrot	1	9
passage	2	AA
pater	1	X
patron	1	17
pencil	1	40
person	1	AA
pheasant	2	3
physics	1	3
pirate	2	12
piston	1	5
platform	1	35
pliers	1	X
pocket	1	A
poolewe	1	
power	1	AA
primrose	1	9
prism	1	5
prophet	3	25
proton	1	
punter	1	
rajah	1	1
rampage	1	X
rangers	3	3
rapist	1	
rector	1	6
regal	1	6
reindeer	1	6
reptile	1	8
research	1	22
rhombus	1	
rhubarb	5	3

Appendix 3 contd.

"List of words generated as Unusual"

words	frequency of generation	Thorndike-Lorge frequency
roman	1	A
romance	1	25
rubber	2	35
rubbish	1	8
ruler	1	32
runner	1	15
sabbath	2	9
safeway	2	
salmon	1	14
sambo	1	
samson	1	4
satan	2	8
saturn	1	7
sausage	14	10
schoolbag	1	
scotland	1	30
seaman	1	5
seesaw	2	
sewage	1	1
sheepdog	1	
shotgun	1	3
sinner	1	8
skeebies	1	
skelpin	1	
skiing	1	
skinhead	3	
skipper	1	7
slater	1	
smoker	1	3
snotter	1	
socket	1	6
solder	2	2
spasm	1	2
spastic	7	
spitfire	1	
stallion	1	1
steamer	1	28
stinker	1	
stoater	1	
stocking	1	34
strainer	1	1
stripper	1	
suedehead	1	
sultan	1	14
support	2	AA
swimming	1	
symbol	1	17
tariff	2	18
teabag	1	

Appendix 3 Contd.

"List of Words Generated as Unusual"

word	frequency of generation	Thorndike-Lorge frequency
tennis	2	18
tiger	1	30
tigger	2	
tombstone	1	2
tonsil	1	3
toucan	1	
trumpet	1	17
tumbler	1	4
turban	1	5
turm	1	1
turnip	1	10
turnup	1	
twilight	1	29
tyrant	1	13
udder	1	2
valour	2	8
vampire	3	1
venture	1	35
verses	1	38
virgin	1	17
waders	1	
warhead	2	
warlock	1	
weather	1	AA
weaver	1	7
werewolf	2	X
wigwam	2	9
window	1	AA
windscreen	2	
wireless	1	11
wizard	6	8
wrangler	1	X
xeon	1	
yashmak	2	
zebra	1	2
zulu	1	1

Appendix 3

"List of words generated as Usual"

word	frequency of generation	Thorndike-Lorge frequency
actor	1	28
aircraft	2	2
aussie	1	
austin	2	3
bacon	1	12
balloon	1	17
barber	1	16
barmaid	1	Y
baton	1	1
bedroom	3	35
belring	1	X
bible	1	25
biro	1	
biscuit	1	14
blackboard	9	9
blazer	4	X
bonnet	1	10
bookcase	3	3
bottle	1	A
breakfast	11	A
briefcase	1	
brother	3	AA
budgie	2	
building	2	AA
butter	2	AA
button	4	39
cabbage	1	16
canal	1	48
candle	1	43
canoe	1	32
canteen	1	2
cardboard	1	5
carpet	1	24
carrot	8	9
celtic	1	2
chemist	3	10
chimney	1	30
chopper	1	2
christmas	1	A
cigar	5	16
classroom	4	3
cobbler	1	2
cockroach	1	1
coffee	1	A
collar	1	44
common	1	AA
concert	1	21
contour	1	4
cooker	1	1
corgi	1	
cornflake	1	

Appendix 3 Contd.

"List of words generated as Usual"

word	frequency of generation	Thorndike-Lorge frequency
corries	1	
council	2	A
country	1	AA
coward	1	22
crombie	3	
cupboard	1	12
curtain	1	A
dentist	4	9
dinner	14	AA
diver	1	4
doctor	5	AA
driver	4	40
duchess	1	12
dustbin	2	
Easter	1	10
elbow	1	26
engine	3	A
evening	3	AA
exam	2	28
exit	1	8
eyeball	1	1
eyebrow	4	12
father	2	AA
fighter	1	14
finger	2	AA
firefly	1	5
floorboard	1	
football	11	26
forenoon	1	9
garden	2	AA
glasses	3	
goldfish	1	1
grandad	1	
grocer	1	11
guitar	2	4
gumboot	1	
gutty	1	
haggis	1	
handle	2	A
headman	1	1
hockey	1	3
homework	1	
horses	1	AA
hotel	1	A
inkstand	1	X
jacket	5	22
jelly	1	19
jesus	1	18
jumper	2	3

Appendix 3 Contd.

"List of words generated as Usual"

word	frequency of generation	Thorndike-Lorge frequency
kettle	2	27
kitchen	4	AA
krispies	1	
lamppost	2	X
leather	1	A
lessons	1	A
lettuce	3	12
levi	2	1
lighter	3	1
lipstick	1	4
madman	1	6
matches	4	A
milestone	1	1
minute	2	AA
mirror	1	46
monday	4	40
money	3	AA
mongrel	1	2
monkey	1	23
morning	9	AA
morris	1	15
mother	3	AA
motor	2	A
moustache	1	1
music	3	AA
needle	2	34
nightime	1	1
nipple	4	X
nostril	1	15
number	1	AA
onion	1	25
oyster	1	23
packet	1	7
panther	1	5
panties	2	
papa	1	34
paper	7	AA
parchment	1	8
parrot	1	9
pavement	4	16
pencil	8	40
pecple	6	AA
pepper	1	27
person	1	AA
picture	4	AA
pilot	2	26
pimple	1	1
pipeline	1	
pocket	1	A
police	2	A
polish	1	28

Appendix 3 Contd.

"List of words generated as Usual"

word	frequency of generation	Thorndike-Lorge frequency
porridge	1	2
portion	1	A
prayer	1	A
princess	1	46
program	3	46
pupil	5	A
railway	1	A
rangers	2	3
record	2	AA
regal	1	6
river	2	AA
roadie	1	
rover	1	15
rubber	3	35
rugby	2	2
rugger	1	
ruler	1	32
sandwich	1	23
sausage	2	10
saviour	1	7
scissors	1	8
seaside	1	4
sheepdog	1	
shopper	1	1
shopping	2	
sidewalk	1	18
singer	1	20
sister	4	AA
skinhead	1	
soldier	2	AA
spaceship	1	
spastic	1	
spider	2	24
spitfire	1	
station	1	AA
steamboat	1	13
student	1	A
subject	2	AA
sugar	3	AA
summer	1	AA
sundown	1	5
sunset	1	24
supper	2	A
tadpole	2	2
taxi	1	17
teabreak	1	
teacher	20	AA
ticket	1	A
toilet	6	11
toothbrush	2	3
tortoise	1	6
traffic	1	36

Appendix 3 Contd.

"List of words generated as Usual"

word	frequency of generation	Thorndike-Lorge frequency
travel	1	AA
triumph	1	41
trousers	7	21
tumbler	1	4
tunnel	1	22
turban	2	5
turnip	2	10
vanguard	1	1
visit	1	AA
washing	1	16
water	1	AA
weather	2	AA
window	9	AA
winter	1	AA
woman	3	AA
woodwork	1	4
woodworm	1	
worker	2	A
workings	1	32
writing	1	39
zebra	1	2
zulu	1	1

Appendix 4

Experiment 2 Tests and Instructions

First, each subject is given a blank sheet of A4 paper. "Could you write your name on the paper - any means of identification will do. Make sure this matches up to that in the second part of the experiment.

1) First of all, I'd like you to write down 10 words. These words must have certain characteristics:
   a) they must be 2-syllable nouns,
   b) they must all be different from one another,
   c) they must be words which you consider unusual, uncommon, infrequently used in the language.

You will have around 15 minutes for this."

After about 5 minutes if the subjects were not making very much progress, the following was suggested:
"Think up different categories of things and then pick an instance of this which is unusual."

2) Having collected the sheets of paper after 15 minutes or before if the subject had finished to his satisfaction, a booklet containing the words to be rated and the vocabulary test was placed face down in front of the subject.
"Again could you put your name on the back of the booklet. Do not turn over yet. There will be a series of words printed on the first 2 pages. I want you to rate each word according to its unusualness or infrequency. That is, beside each word you should put a number - 1 if it is very common, 2, 3, 4, 5 as it becomes more unusual and right up

to 10 which is the most uncommon. So, you have a 10 point scale for unusualness - the same sort of thing you had when you thought up the words. Go through this fairly quickly. Don't stop to ponder a word too long. Put a number beside every word and complete the 2 pages."

All subjects finished within 5 minutes.

3) Vocabulary test

"The next part tries to see if you know the meanings of words, but you get some help since it is a multiple choice answer. Turn over the page.

You have 3 pages of words. The words in capital letters are the ones which you must find the meaning of. You do this by underlining the word from the group of 6 which is most similar in meaning to the word at the top in capitals. Go ahead and complete the 3 pages."

Value 10 is the most unusual; value 1 is the most common.

UMBER

PRINCESS

BATTLE

PAINTING

GHERKIN

FILET

ONION

COPPER

JODHPURS

SCIENCE

CHLORIDE

ANKLET

PIGEON

AMOUNT

PESTLE

PREFECT

BURDEN

TROUBLE

ITEM

SUBSTANCE

WARFARE

LOCKSMITH

CONCEIT

NAPKIN

DWELLING

BARGAIN

SULTAN

PLANET

PASSPORT

FIREFLY

Value 10 is the most unusual; value 1 is the most common

BARREL

COURTSHIP

STEERAGE

FOIBLE

SURTAX

PREVIEW

PRODUCT

MADNESS

PICTURE

DECREE

BOREDOM

APPLE

OCEAN

PERSON

BLOSSOM

BULLET

MONEY

ABBESS

GARRET

CHRISTMAS

FLOWER

DAMSEL

JUSTICE

TABLE

ARRAY

ABODE

WINTER

PASSION

OFFSHOOT

REVOLT

Vocabulary Scale.

In each group of six words below underline the word which means the same as the word in capital letters above the group.

DWINDLE

Swindle pander
diminish wheeze
linger compare

WHIM

complain noise
tonic fancy
wind rush

BOMBASTIC

demoncratic pompous
bickering cautious
destructive anxious

ENVISAGE

contemplate activate
surround estrange
enfeeble regress

GLOWER

extinguish shine
disguise gloat
aerate scowl

LEVITY

parsimony velleity
salutary frivolity
alacrity tariff

AMULET

Savoury jacket
flirtation crest
cameo charm

LAVISH

unaccountable selfish
romantic lawful
extravagant praise

SURMOUNT

mountain descend
overcome concede
appease snub

RECUMBENT

fugitive cumbersome
unwieldy repelling
reclining penitent

TRUMPERY

worthless heraldry
etiquette highest
amusement final

PERPETRATE

appropriate commit
propitiate deface
control pierce

LIBERTINE

missionary rescuer
profligate canard
regicide farrago

QUERULOUS

astringent fearful
petulant curious
inquiring spurious

RUSE

limb	paste
colour	burn
rude	trick

IMMERSE

frequent	hug
reverse	dip
rise	show

VIRILE

demanding	familiar
barbarous	concise
vulgar	robust

STANCE

partition	fixed
position	slope
glance	grief

SENSUAL

controversial	carnal
necessary	crucial
rational	careful

CONCILIATE

congregate	reverse
pacify	radiate
compress	strengthen

LATENT

potential	hostile
overburdened	discharged
ingenious	delayed

CRITERION

superior	critic
certitude	standard
clarion	crisis

FORMIDABLE

tremendous	unexpired
feasible	orderly
ravishing	remembrance

DOCILE

passionate	meek
dominant	homely
careless	dumb

SULTRY

instinctive	severe
sulky	muggy
trivial	solid

EFFACE

rotate	adjoin
disgust	mark
delete	ascend

CONSTRUE

interpret	scatter
contradict	collect
prophesy	anneal

GARRULOUS

ridiculous	daring
massive	ugly
talkative	fast

OBDURATE

formidable	permanent
hesitant	obsolete
exorbitant	stubborn

PALLIATE

regenerate	qualify
alleviate	imitate
stimulate	erase

ADULATE

increase waver
admire prosper
flatter inflate

TEMERITY

impermanence rashness
nervousness stability
punctuality submissiveness

ABNEGATE

contradict decry
renounce execute
belie assemble

VAGARY

vagabond caprice
obscurity vulgarity
evasion fallacy

SEDULOUS

rebellious dilatory
complaisant dilligent
seductive credulous

AMBIT

talisman confines
armature arc
camber ideal

CACHINNATION

guffaw succour
conclave conjunction
cunning controversy

PUTATIVE

punishable computable
supposed worthless
aggressive reconcilable

ADUMBRATE

foreshadow protect
detect eradicate
elaborate approach

FELICITOUS

sincere faithful
valedictory altruistic
voracious opportune

FECUND

esculent optative
profound prolific
sublime salic

TRADUCE

challenge attenuate
suspend establish
misrepresent conclude

SPECIOUS

fallacious coeval
palatial typical
nutritious flexible

NUGATORY

inimitable adamant
sublime contrary
numismatic trifling

RECONDITE

brilliant effervescent
vindictive abstruse
indifferent wise

EXIGUOUS

exhausting prodigous
indigenous esoteric
scanty expedient

MANUMIT

manufacture liberate
enumerate emanate
accomplish permit

MINATORY

implacable diminative
belittling quiescent
depository threatening

Appendix 5

Raw Scores for Experiment 2

Test	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	
Originality of generated words	10	5	9	8	6	20	15	4	10	11	5	0	4	9	2	4	
Fluency	7	5	10	10	5	9	7	5	8	4	5	5	6	6	6	6	
Vocabulary	29	32	25	29	31	29	36	33	29	24	41	27	32	22	26	19	
Rating - m-value words	50	57	51	68	68	52	48	61	65	29	43	53	75	54	73	63	x 100
Rating - Thorndike-L.	50	63	51	46	81	70	68	79	72	73	62	59	76	61	60	71	x 100
Originality Index	36	25	23	20	30	56	54	20	31	69	25	0	17	38	08	17	

	S_{17}	S_{18}	S_{19}	S_{20}	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}	S_{26}	S_{27}	S_{28}	S_{29}	S_{30}	S_{31}	S_{32}	
Originality of generated words	18	10	10	10	9	6	9	7	15	9	9	11	11	5	10	8	
Fluency	10	6	10	9	9	7	7	9	10	6	6	8	7	8	10	9	
Vocabulary	23	31	33	29	21	38	32	31	46	29	29	22	36	24	31	36	
Rating - m-value words	62	48	48	58	73	51	59	53	53	56	54	68	59	24	49	62	x 100
Rating - Thorndike-L.	64	54	67	76	72	71	72	64	69	63	56	66	71	63	68	64	x 100
Originality Index	45	42	25	28	25	22	32	20	38	38	38	35	39	16	25	22	

Appendix 5 (Contd.)

Raw Scores for Experiment 2

Test	Subject Numbers								
	S_{33}	S_{34}	S_{35}	S_{36}	S_{37}	S_{38}	S_{39}	S_{40}	
Originality of generated words	15	13	3	12	14	6	5	10	
Fluency	10	8	10	7	8	9	10	9	
Vocabulary	25	41	23	22	41	22	26	29	
Rating - m-value words	66	48	29	76	61	74	46	50	x 100
Rating - Thorndike-L.	72	52	54	75	65	71	70	72	x 100
Originality Index	38	41	08	43	44	38	13	28	

Appendix 5

Re-analysis of Originality   Experiment 2

Subject Numbers

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	
New Originality *	306	372	656	656	170	270	284	124	372	216	198	344	442	534	458	344	552	x 10
Altered Fluency	6	6	10	10	5	9	7	5	8	4	5	4	6	6	6	6	10	
New Originality Index	510	620	656	656	340	300	406	248	465	540	396	860	737	890	763	573	552	x 10

	S_{18}	S_{19}	S_{20}	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}	S_{26}	S_{27}	S_{28}	S_{29}	S_{30}	S_{31}	S_{32}	S_{33}	S_{34}	
New Originality	232	580	412	524	430	336	628	412	178	208	438	268	622	516	612	498	354	x 10
Altered Fluency	6	10	9	9	7	7	9	10	4	6	8	7	8	10	9	10	8	
New Originality Index	387	580	458	582	614	480	698	412	445	347	548	383	778	516	680	498	443	x 10

	S_{35}	S_{36}	S_{37}	S_{38}	S_{39}	S_{40}	
New Originality	896	278	322	560	732	500	x 10
Altered Fluency	10	6	8	8	10	9	
New Originality Index	896	463	403	700	732	556	x 10

* This score represents the sum of the judges' ratings over the items produced by that subject.

Appendix 6

Experiment 3   Test Material and Instructions

1) Initially each subject was provided with a blank sheet of A4 paper. The instructions were:-

"Could you put your name or initials at the top of the page. You will remember that the last time I asked you to write down 10 words which were (a) two-syllable nouns, (b) were all different from one another and (c) were unusual or uncommon. Well, now I would like you to do the same again but I would like you to try to remember the words you wrote down last time and write them down again. Now, you may only remember say, 4 or 5 words - in this case I would like you to make the total number up to ten, by thinking up 5 or 6 words which are unusual 2-syllable nouns. So you have to write down 10 unusual 2-syllable nouns which are different from one another. Go ahead."

When the subject had finished or 12 or 13 minutes were up: 2) he was presented with the booklet which appears on the following pages. The list consisted of 40 words randomly selected from the words thought up in the previous experiment. The instructions were:

"The next task is again similar to one you have done. I would like you to rate a list of words, 40 in all, for their unusualness. Value 10 is the most unusual, and value 1 is the most common, so you have a ten point scale for each word. You may recognise some of the words as they are taken from the words you thought up the last time. Go ahead."

3) The Wide Range Vocabulary test.

The test is presented in the form shown on the following pages. The instructions were as follows, after filling in name,

"In this test you are to underline the word at the right which will best complete the sentence.
To illustrate: 'A street is a ..... field, hill, road, stream, path.' Which one of these tells what a street is? Road. A line should be drawn under 'road' to show that it is the correct answer. Now do the others in this way. If you are not sure, guess. When you have finished the first page, turn over the test and go right on."

Value 10 is the most unusual; Value 1 is the most common.

VORTEX
BASKET
CABAL
SOFA
COHORT
CODA
SATCHEL
ONYX
TUMULT
ROTA
OGRE
JUMPER
CRITIQUE
CRAMPON
NOGGIN
TALON
STIPEND
ZEBRA
SPITTOON
TABLE

INSTEP
COMMODE
EXCERPT
SALVAGE
HAHA
GLOAMING
RIVAL
CAULDRON
CARRIAGE
MANTEL
PLUTO
CONVERT
RAVINE
NOVEL
POGROM
MANSION
STEAMER
MATER
SATYR
CELLO

WIDE RANGE VOCABULARY TEST

C. R. ATWELL and F. L. WELLS

Form B

NAME.. DATE..

BIRTHDATE.. GROUP..

A.	A **street** is a	field hill road stream path
1.	A **saucer** is a	table spoon hat eat dish
2.	**Jelly** is eaten on	bread potatoes cabbage soup lobsters
3.	To **learn** is to	jump give fall know wake
4.	**Men** are	dogs statues women people monkeys
5.	The **stomach** is for	eating fighting hunting success exercise
6.	If we are **merry** we are	sad married happy drunk naughty
7.	To **step** is to	ride fall stop write walk
8.	We **fry**	cookies eggs coffee people flowers
9.	To be **furious** is to be	angry gentle pretty silly noisy
10.	A **spade** is used to	insult dig rake carry win
11.	**Flutter** refers to	wings drinking singing heels teeth
12.	**Like** means	same different lady new candy
13.	**Bran** comes from	fish peaches wheat bananas liver
14.	**Wealth** is	bananas strength happiness presents riches
15.	A **scholar** is a	fool pendant book student birch
16.	To **agree** is to	argue consent flavor love upset
17.	A **warrant** is served by a	cafeteria preacher restaurant salesman policeman
18.	A **major** is an	artist officer auditor orator igloo
19.	To **preserve** is to	save water fish brown boil
20.	A **cave** is a	rock lake coat hole porch
21.	**Many** means	several mica coins less some
22.	**Spinal** pertains to	fish collarbone architecture backbone disease
23.	To **fidget** is to	scream squirm forget mend rest
24.	To **recognize** is to	talk overlook know ignore seem
25.	**Transact** refers to	business bridges streetcars theaters churches
26.	To **achieve** is to	deceive ravage acknowledge pass accomplish
27.	To **rumple** is to	sit iron dance wrinkle ride
28.	To **take** is to	send please carry lose give
29.	A **zone** is an	acre estate era area antiseptic
30.	A **far** country is	away near beautiful strange rich
31.	**Rickets** is a kind of	medicine disease furniture game food
32.	**Temperature** refers to	electricity dampness pressure heat sunshine
33.	A **couch** is a	cold porch bed chair lie
34.	A **ladle** is a	star crib dipper canoe lady
35.	A **seafarer** is a	captain ship bird reprobate sailor
36.	To **resume** is to	stop continue start consider smoke
37.	**Unfruitful** means	unproductive frosted bitter unfaithful green
38.	To **forewarn** is to	forearm forbear forget forgive foretell
39.	To **whir** is to	eat laugh buzz wiggle cut
40.	**Immune** means	exposed vast diseased inundated protected
41.	To **seclude** is to	travel suspect withdraw linger mistrust
42.	**Rations** refer to	food logic soldiers banks countries
43.	A **coiffure** is a	negligee headdress drink bracelet box
44.	To be **ruthless** is to be	pitiful punishing competitive pitiless aggressive
45.	A **denial** is a	refusal proposal declamation cock confirmation

(over)

46. A **lathe** is a kind of — bath building onion machine clock
47. **Straddle** refers to — babies fighting position money leather
48. **Inquisition** means — punishment war pogrom riot investigation
49. To **relapse** is to — climb recover backslide stop bend
50. A **kingdom** is a — monastery country palace capitol fish
51. To **recruit** is to — discount retreat enlist march fight
52. A **leer** is a kind of — dance beckoning vegetable payment look
53. To make a **pun** is to — laugh rhyme joke fasten kick
54. To **coil** is to — ravel strike wave pin wind
55. A **Calyx** is a term in — physics chemistry orthopedics botany agronomy
56. To **rejuvenate** is to make — young happy beautiful silly blonde
57. To **foil** is to — arrest prevent avoid flavor squeal
58. A **clubfoot** is a kind of — gadder plant society deformity animal
59. A **bilge** belongs to a — wheelbarrow automobile ship tree fish
60. A **flagstone** is used for a — pole weapon sundial tracing pavement
61. To **shroud** is to — bury shiver shape cover worry
62. To be **lenient** is to be — heavy tolerant languorous lithe dependent
63. To **rile** is to — laugh consider anger draw envy
64. To **assent** is to — dissent climb trust fortify agree
65. A **dilemma** is a — problem horn controversy digression contradiction
66. **Infallible** means without — religion error permission science legality
67. A **zigzag** path is — narrow rough up-and-down back-and-forth roundabout
68. **Harum-scarum** means — ambiguous Mohammedan elfish flighty frightened
69. An **azalea** is a kind of — moss fish insect flower chiffon
70. One may **incur** — speed measles spinach people debt
71. To **administer** is to — squander manage substitute judge partake
72. To **exemplify** is to — enlarge exonerate illustrate distrust placate
73. **Manifold** means — many duplicate multiform few simple
74. To **dupe** is to — poison dress deceive demolish clean
75. A **chalice** is a kind of — plate collar cup knight quest
76. A **sot** is — bald neat shiftless stubborn insane
77. To **indict** is to — charge prosecute arrest acquit sentence
78. **Presentiment** means — foreboding gift official emotion chastisement
79. **Avidity** means — greediness dampness dryness hatred honesty
80. **Adjutant** means — bookkeeper officer marine initiation society
81. **Anterior** refers to — back side front right left
82. A **wench** is a — man witch girl nut tool
83. **Malachite** is a kind of — mineral disease race lumber cave
84. To **venture** is to — risk have explore conquer tease
85. A **guise** is a — feature semblance volcano masquerade posture
86. A **tetrasyllable** is a — phrase sentence ruler word quadruped
87. To **inter** means to — debate bury embalm question undertake
88. A **nuncio** is a — pope traveler monastery foreigner messenger
89. A **micrometer** measures — space sound intelligence strength heat
90. **Corvine** means like a — cow hawk crow eagle cat
91. A **mendicant** is a — tailor friar minstrel beggar druggist
92. **Prodigal** is — wasteful masculine thrifty wandering favored
93. A **privilege** is a — kitchen right letter crime favor
94. A **minster** is a — bachelor lady helper lobby church
95. **Phthisis** is a term in — metallurgy astronomy physics psychology medicine
96. An **ibex** is a kind of — bird goat fish jewel plant
97. A **canard** is a — vegetable steamer hoax newspaper fish
98. **Pensile** means — hanging thoughtful written criminal worthless
99. A **spiracle** is for — climbing breathing drawing decoration antisepsis
100. **Eglantine** is a kind of — lily rose violet columbine daisy

Printed in U.S.A. The Psychological Corporation, 304 East 45th Street, New York, N. Y. 10017 71-128T

Value 1 is the most common;

Value 10 is the most unsual.

PARLOUR	ASPIC
KIWI	TUSSOCK
TROTHPLIGHT	SCUTAGE
ZEITGEIST	FORAY
WERGELD	POTTAGE
BUTTER	SAVIOUR
DROPPER	PHYLUM
DOGFISH	SOMETHING
RAJAH	HUMUS
CINDER	TAGGER
SOMA	FACET
GLUTTON	PLUMBING
LEGEND	TASSLE
MIDGET	SPECTRE
PLIMSOLL	BIRO
WINDOW	PIGLET
ACORN	CLANGER
BUTTON	WASTER
PURPLE	RABBIT
PAUPER	CHAOS

Value 1 is the most common;

Value 10 is the most unusual.

TRIPTYCH	DINGO
CHALICE	ANGLER
HAMLET	STURGEON
RECTOR	HANGMAN
SPINDLE	WEEVIL
FEEDBACK	IMPULSE
TECHNIQUE	ETHER
SIDEBOARD	SIDEBURN
CHALET	TUDOR
BODICE	HOVEL
GANTRY	FETISH
HANDCUFFS	CUFFLINKS
PALMIST	SEANCE
BAMBOO	BABOON
PATRE	OCHRE
OGRE	MOUSTACHE
TOUPE	LEGGINGS
CORRIE	ARETE
ELBOW	ROTA
KNEECAP	BUSBY

Value 1 is the most common;

Value 10 is the most unusual.

CONKER	GARAGE
MOTEL	BUMPERS
SATCHEL	ANVIL
CAULDRON	HERMIT
TALON	BRACELET
CANTOR	GIMBAL
QUAGGA	SOZZLE
ILEX	POKER
DINGHY	SNOOKER
GARBAGE	VISCOUNT
PIEBALD	CAMPHOR
TANDEM	ENSIGN
POTAGE	LODGER
PEACOCK	IKON
TADPOLE	BEDSOCK
CHESTNUT	GANDER
LIFELINE	STERNUM
NOSTRIL	BONFIRE
BURETTE	TANGENT
KINGCUP	KAPOK

Value 1 is the most common;

Value 10 is the most unusual.

POLIS
SIBLING
PODZOL
PHLOEM
CUDGEL
PHOTON
MORTAR
HOURI
COFFIN
COALITE
PUNDIT
CANTON
CARTEL
MASON
POGROM
ZEBU
RUMMY
WATTLE
AMBER
DRAGON

MATRIX
DIGIT
XYLEM
BASALT
CUTLASS
STYLUS
HAGGIS
LIMPET
ROUGHAGE
PRATTLE
BANTER
EUNUCH
GNOSTIC
RAMPART
CUDLIP
POGO
HATTER
VELVET
PENGUIN
SKEWER

Appendix 7

Experiment 3: Initial Raw Scores — Subjects (Nos. from exp. 2)

	S_{39}	S_{27}	S_{12}	S_{23}	S_{19}	S_4	S_{21}	S_{36}	S_{31}	S_8
Exp. 2 Originality Index	12.5	37.5	0	32.1	25	20	25	42.8	25	20
Exp. 3 Originality Index	31.2	37.5	66.6	56.2	33.3	25	20	33.3	20	29.1
Mill-Hill Vocabulary	26	29	27	32	33	29	21	22	31	33
Rating of m-value words (x100)	46	54	53	59	48	68	73	76	49	61
Rating of T-L words (x 100)	70	56	59	72	67	46	72	75	68	79
Wide Range Vocabulary	78	81	80	81	89	87	75	77	88	83
Rating of words from Exp.2 (x100)	46	37	32	40	28	37	50	53	52	53
Rating Scale Index (x 100)	60	82	93	83	91	87	83	74	66	79

	S_{14}	S_{13}	S_{34}	S_{16}	S_{11}	S_{26}	S_{28}	S_{37}	S_{30}	S_{35}
Exp. 2 Originality Index	37.5	16.6	40.6	16.6	25	37.5	34.3	43.7	15.6	75
Exp. 3 Originality Index	31.2	46.8	33.3	25	31.2	70	41.6	43.7	32.1	32.1
Mill-Hill Vocabulary	22	32	41	19	41	29	22	41	24	23
Rating of m-value words (x100)	54	75	48	63	43	56	68	61	24	29
Rating of T-L words (x 100)	61	76	52	71	62	63	66	65	63	54
Wide Range Vocabulary	67	82	91	78	88	85	79	86	86	73
Rating of words from Exp. 2 (x 100)	38	65	58	52	31	54	42	40	47	44
Rating Scale Index (x 100)	78	78	75	66	108	87	105	70	62	53

Appendix 7 Contd.

Re-Analysis of Originality based on judges ratings.

	Subjects																			
	39	27	12	23	19	4	21	36	31	8	14	13	34	16	11	26	29	37	30	35
Exp. 2 Originality Index	73	34	86	48	58	65	58	46	51	24	89	73	44	57	39	44	54	40	77	89
Exp. 3 Originality Index	44	43	28	36	42	37	46	39	47	41	64	50	42	52	47	36	28	46	60	54
Rating of words from Exp.2 (x 100)	71	65	58	81	63	71	80	76	87	87	65	65	59	79	71	80	66	69	75	82
Rating Scale Index (x 100)	80	109	125	111	122	117	111	99	88	106	105	105	100	88	145	117	141	93	83	71
Originality Index of Recalled Words	94	34	69	46	60	38	59	30	49	25	92	58	33	53	30	80	59	44	67	93

Appendix 8

"Words generated as Unusual in Experiments 2 and 3"

(Note: blank in 'generated' column means word appeared only once)

Word	Mean judges rating	T-L frequency	Frequency of generation
Aardvark	2.4	1	
Abode	6.6	16	
Agate	3.8	2	2
Alcove	6.6	3	
Alder	3.8	8	
Aloes	3.6	1	
Ampere	3.6	X	
Ampule	3.6		
Anchor	7.6	26	
Annexe	5.2	6	
Arcade	6.6	1	
Argus	2.2	2	
Arras	2.2	1	
Aster	4.2	3	2
Aura	3.8	1	
Balmcake	3.0		
Bangle	6.4	X	
Bantu	3.8	X	
Basil	6.8	1	
Basket	9.6	A	
Basto	1.2		
Bauble	5.0	1	
Bilker	1.6		
Bivalve	2.6	1	
Bogey	4.6	X	
Boiler	8.6	11	
Bolster	6.0	2	
Bootleg	7.2	3	
Bottle	9.6	A	
Bracelet	8.2	10	
Buffoon	3.4	1	
Burette	1.6		
Burgee	1.2		
Byline	3.0		
Bypass	8.4		
Cabal	2.4	1	
Callas	1.2		
Camel	5.0	18	
Cancan	5.2		
Canoe	7.6	32	
Capo	1.6		
Capstan	4.6	X	
Carbar	7.0	23	
Cardboard	9.8	5	
Camage	3.0	1	
Carriage	7.8	46	
Cascade	6.4	5	
Cauldron	6.4	1	2

Word	Mean judges rating	T-L frequency	Frequency of generation
Cavern	8.0	13	2
Cello	5.8		
Chalice	4.2	3	2
Chancel	3.6	1	
Chemise	3.4	3	
Cipher	5.2	3	
Closet	6.4	20	3
Cluster	7.8	23	
Code	3.4	21	
Cohere	3.2	1	
celloid	2.2		2
Comma	7.6	1	
Commode	5.2	X	
Compass	8.0	28	
Concept	7.2	3	
Contract	8.2	A	
Coptic	1.8	X	2
Cordon	6.4	X	
Corrie	2.8		
Coster	2.6	X	
Coven	3.2		
Crampon	3.8		
Cravate	5.4	5	2
Crier	5.6	1	
Critique	4.8	X	
Crofter	5.6		
Cubit	3.4	2	
Cuddy	1.6		
Dactyl	2.2	X	
Dairy	8.6	13	
Damask	3.4	4	
Datum	3.6	1	
Dictum	4.6	1	
Digger	5.8	5	
Digit	6.2	X	2
Diode	1.8		2
Diploid	1.8		
Divot	4.6		
Dodo	5.4		
Dogma	4.2	3	
Domain	5.4	9	
Dragon	6.2	22	
Draper	7.6	X	
Dungeon	6.6	11	
Earwig	7.2		
Emu	5.4		2
Epoch	4.8	5	
Esquire	5.8	10	
Ether	5.8	4	
Excerpt	7.2		
Eyrie	5.0	1	2
Fellow	7.8	AA	
Fennel	4.0	1	
Ferret	6.6	2	
Fever	7.6	45	
Flagon	5.8	2	

Word	Mean judges rating	T-L frequency	Frequency of generation
Flotsam	5.6	X	
Flower	8.8	AA	
Football	9.8	26	
Forelock	4.6	1	
Foyer	6.2	1	
Frigate	4.4	3	
Fructose	3.4		
Furlong	5.6	2	
Gaffer	5.8	X	
Gaiters	4.4	1	2
Galley	5.4	8	
Gangster	7.2	6	
Gargoyle	5.4	X	
Garter	5.6	5	
Gavel	2.6	X	
Genie	3.2	1	
Gesture	8.0	28	
Gewgaw	2.2	X	
Gloaming	3.4	X	
Goggle	5.4	1	
Grampus	2.0		
Gremlin	3.0		
Griffin	4.0	1	
Guitar	8.4	4	
Gunwhale	4.8	2	
Haha	3.8		
Hansel	1.4	1	
Hardware	8.0	8	
Harness	7.6	29	
Harpoon	6.4	1	
Hatchet	6.4	8	2
Haystack	8.6	1	
Hemlock	3.2	7	
Hermit	5.4	16	
Highway	7.4	34	
Hinny	1.0		
Hobnail	5.4	X	
Hoist	3.0	10	
Hooker	5.0	2	
Hotel	7.6	A	
Hovel	6.0	2	
Hummock	3.4	1	
Hydra	3.6	1	
Icon	3.2	X	2
Incest	5.4	X	
Index	7.6	14	
Instep	6.8	1	
Intent	8.4	23	
Isthmus	3.8	13	
Jester	7.2	3	
Jumper	8.8	3	
Jury	8.8	26	
Kayak	4.2		
Kennel	8.4	6	

Word	Mean judges rating	T-L frequency	Frequency of generation
Labrum	1.0		
Lacrosse	4.6	X	
Laggard	4.0	2	
Lagoon	5.0	4	
Lancer	3.4	X	
Landau	2.0	X	
Larynx	5.0	1	
Lasso	4.6	X	
Latrine	5.2	X	
Ledger	6.0	2	
Legal	8.0	28	
Leggings	3.6	3	
Lichen	5.2	7	
Lighthouse	8.2	7	
Limpet	5.2	1	2
Lintel	3.6	1	2
Llama	4.4	1	
Locket	6.4	1	
Logic	6.4	11	2
Lorry	8.6	5	
Lynching	6.6		
Maelstrom	2.6	X	
Magic	7.6	39	
Maiden	6.2	45	2
Mainsail	5.8	1	
Mammoth	5.0	3	
Mansion	8.2	18	
Mantel	4.2	4	2
Marshal	5.4	18	
Mason	6.8	14	
Mater	2.2	1	2
Matins	4.8	2	
Mercer	3.4		
Mirage	6.2	1	
Mistral	3.8		
Mitten	5.4	9	
Moloch	1.0	X	
Moraine	3.0	X	
Motet	3.0		
Mountain	8.4	AA	2
Muck stack	1.2		
Mural	6.2	3	
Musquash	2.8		
Nadir	3.2	X	
Noggin	4.2	Y	
Nomad	2.6	2	
Nova	4.4		
Novel	9.0	39	
Odour	7.2	27	
Ogre	6.8	2	
Onus	5.0	Y	
Onyx	3.2	1	
Optic	7.0	2	
Orbit	7.4	10	
Organ	7.8	48	
Osprey	6.0		2

Word	Mean judges rating	T-L frequency	Frequency of generation
Pallet	6.0		
Panzer	4.8		
Papoose	4.0	X	
Parlour	5.8		2
Passion	8.8	A	
Pathos	3.0	3	2
Pepsin	2.8	X	
Percept	5.0	X	
Phallus	4.6		
Phenol	1.4	Y	2
Phial	4.2	1	
Photon	3.8		
Pilgrim	5.8	14	
Pinball	7.4		
Piton	2.8		
Plaintiff	4.0	3	
Plasma	6.0	1	
Plimsoll	7.0		2
Pluto	4.6	3	
Pogrom	4.4		
Portal	4.2	7	2
Potshed	3.0		
Pottage	2.6	1	
Printer	9.2	7	
Prism	4.4	5	
Process	8.2	A	
Profile	8.0	5	
Profit	9.0	A	
Proton	2.2		
Provost	4.2	1	
Prowler	8.2		
Prudence	7.0	7	
Pulsar	3.0		
Puma	4.6	3	
Pylon	6.6	X	
Quotient	5.0	1	
Rabbit	8.8	43	
Racket	8.4	10	
Ravine	5.6	9	
Respite	5.0	4	
Revolt	9.4	22	
Rhombus	2.2		
Rhythm	8.2	7	
Rickshaw	4.4		
Rival	6.4	32	
River	9.2	AA	
Rota	5.4		
Rowlock	4.6	X	
Rubric	3.4	X	
Rumble	7.4	11	
Sackbut	2.0	X	
Saddle	7.4	41	
Saga	6.6	1	
Sahib	2.4	1	
Salvage	7.0	3	
Satchel	7.6		

Word	Mean judges rating	T-L frequency	Frequency of generation
Satin	7.2	15	
Satyr	2.6	5	
Seaside	9.0	4	
Serfdom	4.6	X	
Settee	8.4	1	
Sextant	4.2	X	3
Shadoof	2.2		2
Shipper	5.2	2	
Sibling	3.4		
Sitar	5.8		
Skivvy	6.4		
Skylab	7.0		
Slipper	9.2	20	
Sneaker	4.0		
Sofa	7.4	14	
Sojourn	5.0	7	
Soma	1.6	Y	
Sonnet	5.2	8	
Sortie	4.6	X	
Sower	3.2		
Spanner	7.6		
Spinney	4.8		
Spittoon	3.8	X	
Spondee	1.0		
Sporran	5.2		
Sputnik	5.8		
Staircase	9.2	9	
Stamen	5.0		
Stanza	4.8		
Steamer	8.2	28	
Stele	1.8	X	
Stipend	3.8		
Strata	7.2	1	
Summit	8.4	17	
Sunrise	9.2	13	
Synapse	2.4		
Syntax	4.6	1	
Syringe	7.6	X	
Table	10.0	AA	
Taboo	7.6	2	
Tallow	4.0	4	
Talon	3.6	2	
Tambone	3.2		
Tandem	7.0	X	
Tantrun	6.8	1	
Tarmac	7.8		
Tatting	3.4		
Teapot	9.4	2	
Tempo	7.2	1	
Tenor	6.8	6	
Terrace	7.8	18	
Thunder	8.8	46	
Tingle	7.0	7	
Torso	6.6	1	
Torus	2.6		

Word	Mean judges rating	T-L frequency	Frequency of generation
Touchline	7.6		
Toupee	7.0		
Trio	7.6	2	
Tropism	3.8		
Trousers	9.0	21	
Tuba	4.4	X	
Tumbril	3.0	X	
Tumult	5.0	11	
Turnpike	5.8	2	
Vampire	7.0	1	
Veneer	5.4	1	2
Verger	5.6	X	2
Victim	8.6	36	
Vortex	3.2	X	2
Walloon	2.6	X	
Wardmoot	1.8		
Wearer	8.4	3	
Wimpole	2.6		
Xenon	1.2		2
Xenxes	1.6	2	2
Xylem	2.2	Y	
Yardang	1.2		
Yeti	3.6		
Yoghurt	8.4		
Zebra	6.2	2	
Zircon	1.6	X	

Appendix 9

Experiment 4: Test Materials and Instructions

Initially each subject was provided with a blank sheet of paper.

"This is the first part of an experiment involving your knowledge of words. This part is done in a group but the second part will be done individually.

Could you please write your name at the top of the page - any means of identification will do, as long as you use the same form of identification from one part of the experiment to another.

First of all, I'd like you to write down English words. These must have certain characteristics:

a) they must be 2-syllable nouns e.g. Bedford College

b) they must all be as different as possible

c) they must be words which you consider unusual, uncommon, infrequently used in the language.

You will have 10 minutes to write down as many words as you possibly can. Go ahead."

After collecting these answer sheets, the Vocabulary test is given out in the form shown previously.

"This part deals with the meaning of words. You get some help from the fact that it is a multiple choice set-up.

Again write your name at the top.

You have 3 pages of words - the words in capital letters you must give the meaning of. You do this by underlining the word from the group of 6 which is most similar in meaning to the word in capitals. If in doubt, guess.

Go ahead and do the three pages."

This session lasted around 30 minutes.

The second session took place a number of weeks later. Again the subject was provided with blank sheets of paper. The instructions were:

"I'm going to ask you to think up some English words again. There will be three different types or groups and they will have different characteristics. Each group will be dealt with separately and I will tell you about each one at the beginning of each section. One general thing is that all the words must be nouns.

You will have 5 minutes for each group and you are to think up as many words as possible. I want you to write down the words as they come to you and immediately you have done this you are to say the word aloud, so that it can be recorded on tape. This is so that I have an idea of when you thought up the word. Are there any questions about this?

a) First of all, I would like you to write down as many words as possible which are a type of container, something which holds something, a receptacle. Go ahead."

Tape recorder was started just before the stopwatch and the signal to the subject.

The subject was stopped after $4\frac{1}{2}$ minutes. The order of the presentation of the conditions was randomised over subjects, so that this is only one possible order.

Now take a new page.

b) "The second category too, have to be nouns. This time write down words which have the sound 'ch' in them, that is ch as a soft sound. This is distinct from the sound in 'chemistry' or 'loch'. As well as having the sound the word must not have it at the beginning. It must not be the first sound. So this would rule out 'church' even though it has the correct sound at the end. Is that clear? Nouns with the sound 'ch' soft but not at the beginning. Go ahead."

Stop after $4\frac{1}{2}$ minutes.

c) "Once again take a fresh page. This time the nouns have to be a certain shape. They must have 2 letters below the line and none above. That is, they must have 2 letters which have a tail, like g, j, y, p, q and no more than 2. And they must not have letters like b, d, l, k, t, etc. (here the shape of the silhouette of such a word is illustrated to the subject). Do you understand? It is not as difficult as it sounds. O.K. go ahead."

Stop after $4\frac{1}{2}$ minutes.

"Good, now the last part is much shorter. Essentially I want you to tell me which words in these categories you know and which you don't know. If you look at the booklet you will see that you have to put a letter beside each of the words on the 3 pages. Put 'A' if you know what the word means and could define it and could use it correctly, put 'B' if the word is familiar but you are not sure of its exact meaning and put 'C' if you have never heard of the word and have no idea of its meaning. Is that clear? Go ahead."

Most subjects finished in 5 minutes. A copy of this booklet is shown on the following pages. The 3 pages are put in different random order for different subjects, though the order of words on any one page remains fixed.

Place either A,B, or C beside each of the words, according to the following scheme: -

A : you know what the word means, could define it and could use it in the correct way.

B : the word is familiar, but you are not sure of its meaning.

C : you have never heard of the word, and have no idea what it means.

HATCHET

PAPER

CISTERN

RESEARCH

APPEARANCE

CRIB

MONOGAMY

JARGON

COPPER

PRIVACY

PEACH

MAGPIE

SALVER

SYRINGE

DRESSER

CANTEEN

PRIMACY

CROCHET

PONCHO

JAG

PITCHER

ENCHANTMENT

PERSONAGE

PAROXYSM

SCRIP

PEP

CUP

JERSEY

CUPBOARD

PARACHUTE

PRESAGE

GRAMERCY

HAUNCH

BUCKET

PIRACY

MOUSETACHE

GRANARY

HORN

UPRISING

BROOCH

ARK

PASSAGE

BUTCHER

GOBLET

CALABASH

INQUIRY

CENSER

MUGWUMP

CREEL

GROUP

COLANDER

DISCHARGE

MANGER

PURVEYOR

PAJAMAS

KITCHEN

PREY

POUCH

PATCH

DITCH

YEOMANRY

ENGINEERING

CYPRESS

PRESS

BUTT

PAGAN

DACHSHUND

CASE

MISCHANCE

KETCHUP

BOOT

INTERCHANGE

REPROACH

RATCHET

VESSEL

BLEACHERS

TUREEN

CROCK

CASTER

ESCUTCHEON

CASK

AVALANCHE

TANKARD

WITCH

TUMBLER

LADLE

SCHEDULE

MACHETE

PARCHMENT

BENCH

Appendix 10

Experiment 4: Raw Scores

	S_1	2	3	4	5	6	7	8	9	10	11	12	13	14
Originality Index $\times 10^{-1}$	4.48	4.91	7.63	6.21	5.28	4.95	5.00	4.81	5.78	5.57	6.41	4.20	4.18	7.19
Fluency	35	7	2	3	11	8	9	17	7	10	9	7	11	19
Vocabulary	20	23	24	24	24	26	27	27	27	32	33	37	39	35
Words Generated - Visual	7	3	8	4	8	9	6	8	4	7	7	6	8	8
Words Generated -Acoustic	8	10	18	9	14	8	6	17	8	14	11	13	9	26
Words Generated -Semantic	21	16	14	21	14	25	15	16	22	28	14	16	19	25
Words Known - Visual	26	18	21	23	20	22	22	24	21	24	26	18	24	27
Words Known - Acoustic	27	26	28	28	28	27	30	28	27	28	28	24	26	29
Words Known - Semantic	25	28	25	26	20	25	27	25	23	28	30	23	25	27

	S_{15}	16	17	18	19	20	21	22	23	24	25	26	27	28
Originality Index $\times 10^{-1}$	8.48	8.12	5.14	6.32	4.53	4.78	5.71	5.16	3.01	7.44	5.47	6.61	5.46	6.65
Fluency	8	13	24	15	4	20	11	15	32	2	12	15	10	9
Vocabulary	43	42	37	36	35	28	28	27	32	34	30	28	31	26
Words Generated - Visual	21	10	6	8	8	12	1	16	10	8	8	8	5	10
Words Generated -Acoustic	32	26	10	9	18	10	8	12	10	16	16	10	5	17
Words Generated -Semantic	42	26	23	22	22	14	16	17	21	17	11	24	26	18
Words Known - Visual	29	27	26	28	25	22	25	20	24	22	21	24	26	22
Words Known - Acoustic	28	29	29	29	30	27	29	30	27	28	30	27	29	26
Words Known - Semantic	30	31	27	27	27	27	28	25	27	24	26	25	28	24

Appendix 11

Experiment 4: List of words generated as Unusual

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Aardvark	4		9	8.23
Acorn	*	10	8	2.63
Adjunct		1	8	7.50
Aeon		X	8	8.25
Agent		40	8	2.50
Airman		X	10	2.60
Aloes		1	8	7.75
Alpha		1	8	6.35
Amber		9	8	4.25
Angler		2	8	5.00
Antics		3	8	5.88
Antique		9	8	3.50
Anvil	2	7	8	4.38
Apex		2	8	5.88
Archives		2	8	4.88
Arete			8	7.75
Aspic		1	8	6.38
Axon			10	7.31
Baboon		2	8	4.88
Bacon		12	8	1.38
Badger		11	8	4.25
Ballet		2	8	2.75
Bamboo		8	8	4.38
Banter		3	8	5.88
Barrel		32	10	1.70
Basalt		X	8	7.50
Bassoon		X	8	6.25
Baton		1	9	5.33
Battle		AA	10	1.30
Bedding		4	8	2.63
Bedsock			8	5.13
Bedstead		2	8	4.88
Biro			8	1.88
Birthday		37	8	2.75
Bison		1	8	6.75
Boater			9	6.67
Bobbin		1	8	4.25
Bodice		2	8	4.88
Bolster		2	8	6.38
Bolus			8	9.38
Bonfire		3	8	3.13
Bowsprit		1	8	6.75
Bracelet	2	10	8	3.13
Brainstorm			8	4.63
Budget		19	9	3.23
Bumpers			8	3.38
Burden		A	8	3.38

* Blank means the word appeared only once.

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Burette			8	8.13
Burgeon			8	8.63
Busby			8	6.13
Buskin		X	8	7.63
Butter	2	AA	8	1.38
Button		39	8	1.38
Camber	2		8	7.13
Camel		18	8	4.38
Camphor		2	8	7.63
Cancer		5	8	2.75
Canter		2	10	4.00
Canton		8	8	6.75
Cantor		X	8	8.00
Caper		7	9	6.75
Caprice	2	2	8	6.38
Capstan		X	8	6.88
Cargo		15	10	2.80
Cartel		X	8	8.25
Cashew		Y	8	4.88
Cauldron		1	8	4.38
Censor		2	8	4.25
Chalet		X	8	4.38
Chalice	2	3	8	7.00
Chaos		9	8	3.88
Cherub		4	8	4.75
Chestnut		16	8	3.38
Cider		5	10	2.10
Cigar		16	8	2.38
Cinder		6	8	4.63
Clanger			8	2.88
Claret		1	8	5.13
Coalite			8	3.88
Cobbler		8	9	4.67
Coccyx	3		8	8.88
Coffin		11	8	3.13
Collage			9	6.45
Comet		15	8	4.75
Condor		X	8	7.13
Conker			8	2.88
Cordon		X	8	4.00
Corrie			8	7.50
Corvette			8	7.75
Cosine	3	X	8	8.13
Country		AA	8	1.63
Cricket		14	8	3.50
Cuckold			8	8.00
Cuckoo		9	8	4.88
Cudgel		4	8	6.13
Cudlip			9	9.38
Cufflinks			8	4.25
Culture		23	8	3.88
Cutlass		2	8	6.25
Cuttle		1	9	8.33

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Dachshund		x	8	4.25
Dada		1	8	7.00
Defile		6	8	5.13
Digit		X	8	4.00
Dinghy		X	8	3.50
Dingo			8	8.63
Discus		X	8	6.00
Doctor	2	AA	8	2.13
Dodo			8	6.75
Dogfish		2	8	4.75
Dogma		3	10	4.60
Dragon		22	8	3.25
Dreadnought		X	8	7.75
Dreamer		4	8	3.75
Dropper			8	4.00
Dyad			10	8.10
Eaglet		X	10	4.90
Easter		10	8	3.00
Effete		X	9	7.78
Ego		X	8	3.25
Elbow	2	26	8	2.75
Engine		A	9	2.00
Ensign		5	8	6.13
Esquire		10	9	4.10
Ether		4	8	6.13
Eunuch		4	8	5.50
Expense			8	2.63
Fable		13	9	2.60
Facet		1	8	4.25
Faggot	2	2	19	4.85
Fancy		A	8	3.00
Fascist		1	10	2.90
Fawcett			8	7.75
Feedback			8	3.13
Feline	2	2	9	6.44
Femur		X	8	6.13
Fetish	2	1	8	3.50
Fetter		8	9	8.22
Finesse		2	8	5.00
Finger	3	AA	9	1.78
Fisher		8	8	3.75
Flipper			9	5.67
Flower		AA	8	1.38
Foetus	3	X	9	4.89
Foolscap		X	8	2.75
Foray		1	8	7.38
Forceps		1	8	6.13
Fortune		A	8	3.25
Fountain		34	8	3.13
Frolic	2	11	10	3.60
Fulcrum		X	8	7.63
Furlong		2	8	6.00

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Gaggle			8	5.63
Gander		2	8	5.75
Gannet			8	6.25
Gantry			8	7.88
Garage		14	8	1.88
Garbage		4	8	3.13
Gibbet		2	8	7.63
Gibbon		1	8	5.38
Gimbal			8	8.00
Giraffe		1	8	3.25
Glasses			8	1.63
Glottis		X	8	7.13
Glutton		2	8	4.00
Gnostic		X	8	7.88
Gobble		5	10	3.80
Goblin		14	8	4.13
Gopher		2	8	7.75
Guitar	2	4	8	2.88
Gumption		X	8	4.63
Guppy			8	8.13
Haggis			8	6.13
Hamlet		15	8	5.13
Handcuffs		1	8	4.75
Handle		A	8	1.50
Hangman		1	8	5.88
Harness		29	9	5.11
Harpy		1	10	6.70
Hashish			10	3.10
Hassle			8	3.00
Hatter		1	8	5.63
Heather		5	8	3.88
Helix			9	7.56
Hermit		16	8	4.38
Hillock		3	9	5.22
Hobo		1	8	6.75
Horseshoe		4	9	3.89
Houri		X	8	9.25
Hovel		2	8	3.63
Humus		2	8	5.13
Husky		7	8	5.38
Hybrid		5	9	6.33
Iceberg		4	10	2.90
Ikon		X	8	7.63
Ilex		X	8	9.50
Import		31	9	3.56
Impulse		26	8	3.00
Index		14	9	3.56
Ingot		1	8	7.88
Islet		3	8	7.00
Jacket		22	8	1.88
Judgement		A	8	4.50

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Kapok		x	8	7.38
Kingcup			8	7.50
Kiwi		X	8	6.00
Klaxon			8	5.75
Kneecap			8	2.88
Lady		AA	8	1.50
Lamprey	2	X	8	7.75
Landau		X	9	9.44
Legend		22	8	3.63
Leggings		3	8	6.63
Letter		AA	8	1.63
Lido			9	7.44
Lifeline			8	4.00
Lighter		1	8	2.63
Lily		33	8	5.13
Limpet	2	1	8	6.00
Locksmith		X	9	5.56
Lodger		1	8	3.25
Logic	2	11	9	3.67
Lotus		2	9	6.33
Lyric		6	8	4.13
Manger		7	9	5.11
Mango		1	9	4.89
Mantle	2	19	9	7.13
Mantra	2		8	8.13
Marker		2	8	2.00
Mason		14	8	4.75
Matrix		X	8	4.75
Mentor		1	8	8.88
Midget		2	8	3.88
Miller		8	8	5.38
Mirror		46	8	1.88
Mistress		34	8	3.63
Mitten		9	9	6.33
Mobile		2	8	3.75
Monad		Y	10	8.10
Mongoose			8	7.38
Mortar		8	8	5.75
Motel			8	2.88
Moustache		1	8	4.13
Mystic		4	10	4.70
Navel		X	8	4.38
Neon			9	5.56
Nipple		X	8	2.50
Nodule		1	8	6.13
Nomad		2	8	4.50
Nonet			10	9.50
Nostril		15	8	3.63
Nova			8	4.63
Nuance		X	9	6.67
Number	2	AA	8	1.25

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Occult		2	9	5.22
Ochre		1	8	6.50
Octet			10	5.70
Oddment			10	2.30
Ogive	2		8	9.00
Ogre		2	8	6.50
Onus		Y	10	5.90
Orbit			8	5.50
Ouija			8	9.25
Outcrop		1	9	5.11
Palmist	2	X	8	5.13
Pamphlet		9	8	3.63
Paper		AA	8	1.25
Parlour		30	8	4.25
Party		AA	10	1.10
Passer		1	10	3.90
Pater		X	5	5.80
Pathway		7	8	2.75
Patter		8	8	3.63
Pauper		2	8	4.13
Peacock		10	8	4.13
Pelvis		X	8	4.38
Pencil		40	8	1.25
Pendant		2	8	4.00
Pestle	2	2	8	7.75
Phloem			8	8.25
Phoenix		1	8	4.75
Photon			8	8.13
Phylum			8	7.13
Pickaxe		1	8	5.00
Piebald	2	X	8	5.25
Piglet			8	3.25
Pincer		1	8	4.75
Pistache			8	7.75
Pistol		20	8	4.13
Placard		3	8	4.25
Planet		34	8	3.75
Plankton		X	8	5.50
Plantain		2	8	7.75
Plastic		3	8	2.00
Platoon		1	8	5.75
Pleura		X	9	9.56
Plimsoll			8	3.50
Plumbing		2	8	3.63
Podzol			8	9.38
Pogo			8	7.13
Pogrom			8	7.50
Poker			8	3.38
Polis			8	9.13
Pontoon		X	8	4.63
Porter		19	8	3.00
Portrait		19	8	3.38
Potage			8	6.75
Pottage		1	8	7.00
Prattle		1	8	5.38
Prelate		4	8	7.50
Primate		X	9	5.33

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Prophet		25	8	4.50
Proton	2		8	7.63
Pumpkin		13	9	4.89
Pundit			8	5.38
Purport		4	9	7.11
Purple		37	8	2.25
Quagga			8	9.63
Quasar	2		10	7.80
Quaver		3	9	5.78
Quayside			8	5.25
Rabbit		43	8	2.13
Rabies		1	8	4.00
Rajah		1	8	6.25
Rampart	2	4	8	5.88
Rebate		1	8	3.38
Rebirth		X	8	3.25
Rector		6	8	4.75
Rhombus			10	9.00
Rocket		4	8	4.50
Rota			8	4.13
Rotor			8	4.00
Roughage		X	8	5.38
Rubber		35	8	2.25
Rummy		X	8	4.88
Sable		9	8	8.00
Sabre		4	8	5.00
Saline		1	9	5.67
Samite		X	8	9.38
Sampler		1	8	5.50
Satchel		3	8	3.25
Saviour		7	8	4.75
Scalple		X	9	4.78
Scarab		X	10	7.90
Scribble		3	8	2.63
Sculpture		12	8	4.00
Scutage			8	9.13
Seance		Y	8	4.75
Seaside	2	4	8	2.63
Seizure		3	9	5.56
Sepal		1	9	8.56
Shampoo		1	8	2.75
Sherbet		2	10	2.70
Shipwright		X	9	6.67
Shoulder		AA	8	1.88
Sibling			8	6.00
Sideboard		3	8	2.25
Sideburn	2		8	5.75
Silver		AA	8	2.63
Sister		AA	8	1.63
Skewer		1	8	4.38
Skittle			18	4.72
Smoker		3	8	1.63
Snooker			8	3.88
Snorkle			8	7.00
Solstice		X	8	6.88

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Soma		Y	8	7.75
Something		AA	8	1.50
Sozzle			8	7.25
Spaceship			8	5.00
Spectre		5	8	5.00
Spindle		9	8	5.63
Spittle		Y	8	6.00
Squirrel		24	9	3.56
Stitching			8	3.63
Stocking		34	10	2.00
Stomach		30	8	1.88
Strata		1	9	6.11
Sturgeon		X	8	6.38
Stylus	2	X	16	4.06
Sucrose		Y	8	5.25
Sweater		8	8	2.00
Syndrome			8	5.25
Syntax	2	1	8	5.25
Syringe		X	8	4.50
Table		AA	8	1.38
Tadpole		2	8	3.50
Tagger			8	8.88
Talmud		X	8	8.38
Talon		2	8	5.50
Tampon			10	3.20
Tandem	2	X	16	5.19
Tangent		1	8	5.00
Tapir		X	8	7.88
Tassle (el)		5	8	4.88
Technique		6	8	2.38
Tempo		1	8	3.63
Termite		1	9	7.11
Terrain		X	8	6.38
Tetrarch			10	8.50
Thistle		8	8	5.00
Throstle		X	8	7.75
Tiffin		X	8	7.25
Tiger		30	8	4.38
Tinsel		2	9	5.78
Toadstool		1	10	4.50
Toggle		X	9	7.22
Torrent		14	8	5.00
Toupe			8	6.25
Towel		18	8	1.38
Traction			8	6.63
Triad		X	10	5.30
Tripod		2	8	4.75
Triptych		X	8	9.00
Trothplight			8	9.25
Trumpet	3	17	8	3.00
Tudor		1	8	5.63
Tureen		X	9	7.22
Tussock		X	8	7.75
Twosome		X	8	4.25

Word	Frequency of Appearance	Thorndike-Lorge	No. of Judges	Rating
Ulcer		1	9	4.00
Ulna		X	9	8.44
Umbra		Y	8	8.38
Varnish		6	8	4.00
Vector		Y	9	7.22
Velvet		32	8	2.63
Viper		5	9	5.22
Virus			10	3.00
Viscount		3	8	4.88
Vittel			8	9.00
Vixen		X	8	5.50
Warble		6	8	6.63
Waster			8	3.25
Wattle		X	8	6.75
Weevil		2	8	5.63
Wergold			8	5.63
Wether		1	8	6.63
Window	2	AA	16	1.38
Witchcraft		5	10	3.30
Writer		A	9	2.38
Xylem		Y	8	8.13
Yoga	2		8	3.00
Yoghurt			8	2.13
Zebra		2	8	6.63
Zebu			8	9.88
Zeitgeist			8	9.38
Zero		11	8	2.38
Zymase			8	8.38

Appendix 12 and 17

"Appendix 12 and Appendix 17. Experiments 4 and 6. Words generated in each Category."

(a) VISUAL CATEGORY

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
agency	32		
aggression	3		
agony	24		
appearance	A		
angiosperm	Y		
asparagus	6		
augury	X		
cogency	X		
copper	46		
copy	A		
cosmogony	X		
cypress	6		
egg	AA	4	3
emergency	19		
energy	41		
engineering	15		
engraving	3		
equerry	Y		
espionage	1		
expiry	Y		
gag	3	5	3
gamey	-	1	
gang	25		
gangrene	1	3	1
gap	17	12	2
gape	7	5	
garage	14	3	
gasp	29	3	
gauge	5	5	1
gawp	-	2	
gig	3	5	1
gimp	X		
ginger	13	1	1
gong	4	5	
gorge	9	1	
gossip	16	1	
gramercy	2		
granary	2		
grange	5	1	
granny	17	1	
granpa	9		
grape	34	1	1
grasp	36	3	
gravy	8		
greenery	X		
gregariousness	-	1	
grip	35	2	1
grocery	11		

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
grog	X	1	
grope	10	1	
group	AA		
guy	25	2	1
gym	-		2
gyro	-	1	
imagery	1		
ignominy	1		
injury	24		
inquiry	22		
jag	1	2	
Japan	49		
jape	X	1	
jargon	1	1	
jay	15	7	1
jeep	-	4	1
jemmy	-	1	
jerry	-		
jersey	21		
jig	4	4	
jigsaw	X	2	
jog	-	2	
journey	A		
joy	AA	4	2
jug	8	7	2
jump	A	5	
jumper	3	3	2
juniper	3	2	
jury	26		
magpie	2		
misgiving	5		
monogamy	X		
mugger	-	1	1
mugwump	X		
myopia	X		
napery	X		
nigger	2		
nincompoop	X		
occupancy	1		
opening	36	1	
oppression	8		
orgy	4		
osprey	1		
oxygen	25		
pagan	7		
page	AA	7	2
pajamas	6		
pang	12	3	
pansy	7	1	
paper	AA	5	1
paragon	1		
parsnip	1		
paroxysm	1		

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
passage	AA		
paving	14	1	
pay	AA	13	4
payee	X	1	
peep	27	2	2
peerage	2	2	
peg	9	5	1
penguin	2	1	
penny	38	2	
pep	3	1	
pepsin	X		
personage	8		
pig	44	7	4
pigeon	34	1	
pimp	X	1	
ping	X	2	
pip	X	4	
pipe	A	3	4
pique	6	1	
piracy	1		
poesy	1		
pompion	X		
pompon	X		
pong	-	3	
pongee	X		
pony	32	2	
poop	2		
pop	21	3	
pope	26	2	1
porpoise	1		
porringer	X		
posy	1		
prang	-		
prayer	A	3	
precipice	8		
presage	2		
prey	28	1	
prig	X	1	
primacy	1		
priory	3		
privacy	6		
privy	4		
processionary	1		
programme	46		
prop	9		
prong	1	2	
pug	1	2	1
pump	29		
pup	6	2	
purpose	AA		1
purveyor	2		
pussy	11	2	
pyre	X	1	
pyx	X		

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
quagmire	X		
quarry	11		
quay	3	3	1
query	9	2	1
quip	1	6	
rejoicing	10		
repugnance	1		
sapper	X	1	
saying	13	1	
signory	X		
snapper	X		
sponge			1
sporangia	X		
spray	22	1	
sprig	13		
spring	AA		
spy	30		
supper	A		
suppression	4		
supremacy	9		
surgery	7		
swagger	3		
symposium	1		
synonym	2		
synopsis	1		
syringe	X		
syrup	6		
uprising	6		
urgency	1		
vagrancy	X		
voyage(r)	2		
yap	2	1	
yearning		1	
yeomanry	1		
yogi	1	1	
yoyo			1
zigzag	7		
zipper	-	1	

Appendix 12 and 17 (Contd.)

(b) ACOUSTIC

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
achievement	23	1	
anchovy	1	2	1
apache	1		
approach	A	1	
arch	34	4	2
archer	11	2	
archery	2		
archipelago	3		
archway	1	2	
armchair	6	1	
artichoke	1		
attachment	8		
avalanche	4		
bachelor	13		
batch	3	10	6
beach	A	2	2
beech	10	2	
belch	1	1	
bench	46	3	1
birch	16	3	2
bitch	X	4	6
bleach	9	2	
botch	X	2	
branch	AA	6	3
breach	10	1	
breeches	8	1	
broach	3	1	
brooch	4		
bunch	32	6	
butcher	22	2	1
cache	1	3	
cachet	X	1	
catch	AA	14	6
catchment	-	1	
cinch	1	1	
cinchona	X		
cliche	Y	1	
clinch	3		1
clutch	21	4	
coach	42	2	
conch	X	2	1
couch	28	1	
creche	-	1	
crochet	5		
crotchet	X	1	
crunch	2	5	3
crutch	10	4	1
dachshund	X	1	
debauchery	1		
despatch	1		
detachment	8	1	
discharge	30	1	

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
ditch	28	6	8
duchess	12	1	
echelon	7	2	
enchantment	8	1	
encroachment	2		
escutcheon	1	1	
etching	2		
exchange	A		
exchequer	3		
finch	1	1	1
franchise	4		
fuchsia	X		
gulch	3		
handkerchief	35		
hatch	19	5	4
hatchery	X	2	
hatchet	8	3	1
haunch	3		
hitch	9	2	6
hooch	Y		
hunch	5	3	2
hutch	X	3	2
impeachment	2	1	
inch	AA	1	2
interchange	5	1	
itch	5	5	3
ketch	1	2	
ketchup	X	3	1
kitch	-		
kitchen	AA	3	1
larch	2	6	2
latch	8	15	7
latchet	X	1	
launch	18		1
lecher	-	2	1
leech	4	6	
lichen	7	3	2
loach	Y		
lunch	39	8	6
lurch	3	4	3
lynching	5	1	
machete	Y	1	
machine	AA	2	
manchet	X		
maraschino	1		
march	AA	4	2
marcher	X	2	
match	A	17	6
merchant	A		
merchandise	14		
mischance	2		
mischief	20		
moustache	1	1	
mulch	1		
munch	4	2	
niche	4	6	

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
notch	11	1	
orchard	20	1	
ostrich	5		
panache	-	1	
parachute	5	1	
parchment	8	1	
pastiche	-	1	
patch	34	13	3
paunch	1		
peach	29	4	1
perch	23	2	2
pinch	20	2	3
pistachio	X		
pitch	43	5	5
pitcher	20	2	
poacher	2	3	
poncho	1	1	
pooch	1		1
porch	A	4	
pouch	8	1	
preacher	16	1	
punch	17	1	3
ranch	20	5	4
ratchet	X	2	
reach	AA	4	2
research	22		
reproach	17		
riches	20		2
roach	1	4	
sachet	X	2	
sandwich	23	1	
satchel	3	8	1
schedule	14		
scorch	9	1	
scotch	16	1	
scratch	30	5	1
screech	8	2	
scrunch	-	1	
search	A	5	2
seneschal	2		
sketch	25		
slouch	3		
snatch	27		
speech	AA	1	
spinach	8	1	1
stanchion	1		
starch	10	4	1
stench	1	1	1
stitch	15	4	1
stretch	A		
switch	15	2	
teacher	AA	5	
tench	-	1	
thatch	6	6	1
torch	17	3	
touch	AA	4	4

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
treachery	10		
trench	16		1
trencher	2		
truncheon	1		
twitch	9		
urchin			
vetch	X	1	
voucher	X	1	
watch	AA	11	2
wench	4		1
winch	1	2	4
witch	24	3	2
wrench	11		
wretch	12	3	

Appendix 12 and 17 (Contd.)

(c) SEMANTIC

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
amphora	Y		
ark	7		
ashtray	-	4	
bag	AA	21	13
balloon	17		
barrel	32	14	1
basin	25	8	6
basket	A	12	6
bassinet	X		
bath	46	10	3
beaker	1	6	3
bin	6	12	9
boiler	11		
boot	37	1	
bottle	A	20	9
bowl	A	20	9
box	AA	25	12
breaker	6		
briefcase	-	5	
bucket	16	11	5
bunker	-		1
bureau	44	1	
butt	10		1
cabinet	26	2	
caddy	X		
caisson	1		
calabash	X		
can	AA	10	4
canister	X	1	2
cannikin	X		
canteen	2		
capsule	2		
carafe	10		
carpetbag	X	1	
carton	3	6	3
case	AA	5	11
cask	5	2	
casket	5	2	
casserole	3	2	
caster	X		
cauldron	5	1	
cellar	32		1
censer	2		
chalice	3	1	
chest	41	6	
churn	4	1	
cistern	3	1	
closet	20		
coffer	3	1	
coffin	11	1	1
colander	X	1	1
commode	X		

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
compact	13		
cornet	2	1	
cornucopia	-		1
cot	15	2	
cradle	21		
crate	7	3	1
creel	X		
crib	6		
crock	2		
crucible	2	1	
cruet	X		
cruse	X		
cup	AA	24	11
cupboard	12	8	6
decanter	X	2	
demijohn	A		
dish	A	10	2
drawer	20	10	5
dresser	7		
drum	40	1	1
dustbin	-	4	2
eggcup	-	2	1
envelope	22	4	
ewer	3		
file	43	1	1
flagon	2		
flask	4	9	3
flowerpot	1	1	
folder	2		1
frail	X		
glass	AA	16	6
goblet	7	1	1
gourd	2		
hamper	8	2	
handbag	1	8	5
haversack	X	1	
hold	AA	1	
holdall	-	3	
hopper	2	1	
horn	A		
jar	43	10	6
jug	8	17	5
keg	3		
kettle	27	2	1
ladle	2		
laver	X		
magnum	X		
manger	7		
mug	5	15	5
noggin	Y		
packet	7	6	5
pail	16	3	
pan	A	7	2
pannier	X		
phial	1	1	2
pitcher	20	2	1
plate	A		

Word	T-L frequency	frequency of generation Exp. 4	Exp. 6
platter	8	1	
pocket	A	6	2
poke	12		
pool	34	3	
portfolio	2		
portmanteau	1		
pot	47	12	5
pouch	8	2	
puncheon	1		
purse	38	5	5
quiver	27		
reservoir	8	1	
retort	15		1
rucksack	-	1	2
sack	30	9	2
saddlebag	1		
safe	AA	2	
salver	X	1	
satchel	3	5	
saucer	7	2	
saucepan	7	6	
scabbard	4		
scrip	1		
scuttle	5		
sheath	9		
shell	A	1	
silo	X	1	
sink	A	6	2
siphon	-		1
skillet	2	1	
spittoon	X		
stein	2		
suitcase	5	13	6
tank	19	2	2
tankard	1		
teapot	2	1	4
tender	A	1	
testtube	1	3	2
till	-		1
tin	36	17	12
toby	1		
tray	17	2	1
trough	8	4	
trunk	48	7	5
tub	16	2	1
tube	32	2	3
tumbler	4	2	
tun	X		
tureen	X		
urn	8	1	2
valise	2		
vase	12	14	5
vat	5	2	3
vessel	A	5	3
vial	2		1
wallet	6	3	3
wardrobe	10	3	4
wheelbarrow	2	1	

Appendix 13

Experiment 5: Instructions to Subjects

1. Given to all subjects.

"The experiment is a simple one. You will be presented with some words on this tachistoscope (explained for non-psychologists), and your task will be quite simply to tell me what these words are. There are no tricks involved: no highly emotional words, or very unusual words. What I am interested in solely is your recognition time. This is the time it takes you to identify the word. To do this there is a voice key, a microphone, here (shown to the subject under the aperture) which stops a timer. So all you have to do is to say the word aloud as soon as it appears."

(any questions were now answered).

"Now I am going to give you a number of practice trials in order to get you used to the task and also to establish your threshold for recognition - the fastest presentation at which you can recognise the word. After we do this, all the words in the experiment will be presented well above your threshold in order that you can read them quite clearly. The procedure will be that I will say 'right' and the word will appear in the next couple of seconds."

2. Experimental subjects were given the following additional instructions, after the practice trials.

"Now the words you are going to see will fall into 3 categories and these are:

(a) visual category - all the words will have 2 letters below the line and none above. That is, letters like 'g, j, y' etc., but not 'b, d, k' etc. This will give you some idea of the shape of the word.

(b) acoustic category - all the words will have a 'ch' sound as a soft sound, but it will not appear at the beginning of the word.

(c) semantic category - all the words will mean or be, a type of container, a receptacle, something which holds something else.

Each of these will appear in groups of 5 words, so that 5 cards, one after another, will come from the same category. This will be simplified by the fact that I will tell you which category is going to appear next. So all you have to do as before, is to say the word aloud as soon as you recognise it."

All subjects were given a break of around 2 minutes after 30 presentations.

Experiment 5: Example of the words presented to the subject.

	Word	Previously Generated	Recognition Time (seconds)
AC.			
	cachet	yes	.99
	bitch	yes	.67
	latch	yes	.59
	lunch	yes	.78
	wretch	no	.45
SEM.			
	casket	no	.87
	purse	yes	.76
	carton	yes	.57
	kettle	no	.49
	tankard	no	.67
VIS.			
	jig	yes	.68
	peep	no	.56
	gimp	yes	.78
	grape	no	.71
	jargon	no	.43
VIS.			
	yoga	no	.65
	jigsaw	no	.66
	quarry	no	.73
	yap	yes	.72
	gap	yes	.58
SEM.			
	suitcase	yes	.62
	mug	yes	.62
	jar	no	.64
	churn	no	.90
	envelope	yes	.81
AC.			
	botch	yes	.84
	stench	no	.83
	parachute	no	.73
	ketchup	no	.64
	ditch	yes	.79

	Word	Previously Generated?	Recognition Time (seconds)
AC.			
	stitch	yes	.64
	match	yes	.56
	broach	no	.58
	touch	no	.78
	screech	yes	.71
VIS.			
	jury	no	.72
	pig	yes	.83
	magpie	no	.76
	quay	no	.75
	gang	no	.64
SEM.			
	saddlebag	no	.74
	reservoir	no	.62
	handbag	yes	.63
	urn	no	.61
	bowl	yes	.50
SEM.			
	bag	yes	.51
	packet	yes	.52
	boot	no	.54
	desk	no	.48
	carpetbag	yes	.39
VIS.			
	pug	yes	.79
	pip	no	.76
	penguin	no	.73
	gig	yes	.65
	mugger	yes	.62
AC.			
	lurch	no	.81
	research	no	.80
	wrench	no	.50
	couch	no	.49
	satchel	yes	.60

Appendix 14

Experiment 5 Mean Recognition Times and Set Sizes for 3 Categories

Recognition Times	Experimental Group										
Generated Words	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Visual	70	59	65	69	61	78	56	62	91	60	x 10^{-2} seconds
Acoustic	64	56	62	68	64	75	44	69	74	57	"
Semantic	59	54	53	65	58	74	47	60	77	51	"
Non-Generated Words											
Visual	57	73	61	76	64	83	55	66	95	62	"
Acoustic	60	62	68	73	73	80	52	74	96	60	"
Semantic	57	61	63	71	61	71	51	69	91	65	"
Estimated set size											
Visual	189	158	177	177	183	183	183	170	177	151	no. of words
Acoustic	182	170	182	182	182	182	176	176	176	164	"
Semantic	172	149	166	172	155	160	155	166	149	144	"

Comparison of Experimental Group and Control Group
Mean Recognition Times

Experimental Group	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Visual	62	69	63	73	63	80	55	65	90	61	x 10^{-2} seconds
Acoustic	62	60	66	70	68	77	48	71	88	59	"
Semantic	58	57	58	68	59	73	49	65	75	58	"
Control Group	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}	
Visual	53	57	72	69	53	58	62	66	59	64	"
Acoustic	55	61	80	73	54	52	62	62	59	61	"
Semantic	54	58	67	73	52	55	63	61	57	60	"

Appendix 14

Experiment 5   Data grouped by T-L frequency of occurrence 1

Previously Generated Words

Visual Code	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Frequency - 26 per mill Recogn. time	49	55	58	61	60	72	56	57	76	56	x 10^{-2} seconds
Set size	34	34	34	34	34	34	34	34	34	34	
No. words generated	1	1	3	2	4	5	4	4	1	2	
Frequency - 11-25 p.m. Recogn. time	54	58	60	68	76	85	60	$\hat{71}$	94	67	x 10^{-2} seconds
Set size	25	25	25	25	25	25	25	25	25	25	
No. words generated	1	3	3	2	3	2	6	0	1	3	
Frequency - less than 10p.m. Recogn. time	78	$\hat{65}$	79	75	55	79	48	60	99	57	x 10^{-2} seconds
Set size	108	81	97	97	102	89	108	90	91	78	
No. words generated	6	0	5	3	1	9	3	2	4	4	

$\wedge$ - estimates as no value available

Previously Generated Words (Contd.)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Acoustic Code											
Frequency - more than 26p.m. Recogn. time	57	54	57	69	68	73	45	59	78	57	x 10^{-2} seconds
Set size	33	33	33	33	33	33	33	33	33	33	
No. words generated	9	4	7	3	9	4	6	7	3	4	
Frequency - 11-25 p.m. Recogn. time	64	62	64	69	49	66	45	$\hat{61}$	69	53	x 10^{-2} seconds
Set size	35	35	35	35	35	35	35	35	35	35	
No. words generated	2	2	3	2	2	3	3	0	1	3	
Frequency - less than 10p.m. Recogn. time	68	60	68	67	65	78	42	84	71	63	x 10^{-2} seconds
Set size	94	78	91	91	94	97	81	88	88	81	
No. words generated	14	1	6	6	13	3	6	5	2	1	

^ - estimates as no value available

Previously Generated Words (Contd.)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Semantic Code											
Frequency - more than 26 p.m. Recogn. time	61	54	55	67	57	58	44	56	75	50	x 10^{-2} seconds
Set size	43	43	43	33	43	43	43	43	43	43	
No. words generated	11	10	11	7	5	7	12	14	11	13	
Frequency - 11-25 p.m. Recogn. time	58	51	45	58	55	105	47	52	101	51	x 10^{-2} seconds
Set size	21	21	21	21	21	21	21	21	21	21	
No. words generated	3	2	1	3	3	2	1	4	1	6	
Frequency - less than 10p.m. Recogn. time	57	58	54	67	61	78	51	67	66	54	x 10^{-2} seconds
Set size	96	72	92	94	81	79	74	88	79	72	
No. words generated	10	6	10	4	13	8	6	8	3	6	

Appendix 14

Experiment 5   Data grouped by T-L frequency of occurrence 2

Previously Non-Generated Words

Visual Code	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Frequency - 26 per mill Recogn. time	56	62	57	61	63	78	^ 49	63	99	^ 62	x 10^{-2} seconds
Set size	34	34	34	34	34	34	34	34	34	34	
No. words generated	1	1	3	2	4	5	4	4	1	2	
Frequency - 11-25 p.m. Recogn. time	61	73	73	74	68	78	40	58	71	57	x 10^{-2} seconds
Set size	25	25	25	25	25	25	25	25	25	25	
No. words generated	1	3	3	2	3	2	6	0	1	3	
Frequency - less than 10 p.m. Recogn. time	57	78	53	81	65	91	55	70	83	62	x 10^{-2} seconds
Set size	108	81	97	97	102	89	108	90	91	78	
No. words generated	6	0	5	3	1	9	3	2	4	4	

^ - estimates as no value available

Previously Non-Generated Words (Contd.)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
Acoustic Code											
Frequency - more than 26 p.m. Recogn. time	51	62	^ 58	68	^ 58	67	48	^ 61	72	50	x 10^{-2} seconds
Set size	33	33	33	33	33	33	33	33	33	33	
No. words generated	9	4	7	3	9	4	6	7	3	4	
Frequency - 11-25 p.m. Recogn. time	62	51	82	72	72	76	52	66	79	56	x 10^{-2} seconds
Set size	35	35	35	35	35	35	35	35	35	35	
No. words generated	2	2	3	2	2	3	3	0	1	3	
Frequency - less than 10 p.m. Recogn. time	60	63	62	73	74	77	51	76	100	62	x 10^{-2} seconds
Set size	94	78	99	91	94	97	81	88	88	81	
No. words generated	14	1	6	6	13	3	6	5	2	1	

^ - estimates as no value available

Previously Non-Generated Word (Contd.)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	
<u>Semantic Code</u>											
Frequency - more than 26 p.m. Recogn. time	54	55	65	$\hat{67}$	58	71	47	59	90	57	x 10^{-2} seconds
Set size	43	43	43	43	43	43	43	43	43	43	
No. words generated	11	10	11	7	5	7	12	14	11	13	
Frequency - 11-25 p.m. Recogn. time	$\hat{61}$	49	55	$\hat{65}$	65	63	55	89	84	80	x 10^{-2} seconds
Set size	21	21	21	21	21	21	21	21	21	21	
No. words generated	3	2	1	3	3	2	1	4	1	6	
Frequency - less than 10 p.m. Recogn. time	60	64	62	75	62	73	51	70	79	64	x 10^{-2} seconds
Set size	96	72	92	94	81	79	74	88	79	72	
No. words generated	10	6	0	4	13	8	6	8	3	6	

$\wedge$ - estimates as no value available

Appendix 15

Experiment 6: Instructions to Subjects

"This experiment is in 2 main parts - firstly you will think up some words and secondly I will show you some words and ask you to decide whether they belong to certain categories. Could you put your name at the top of the page."

The normal instructions for word generation followed, as detailed in the appendix for Experiment 5 (13). Three minutes were substituted as the stated time limit for generation.

In the second section the subject spoke his response into a microphone which acted as a voice key and stopped the timer. The instructions were as follows:

"This part involves you in looking at some words, and telling me whether they belong to the categories which you have just dealt with. This machine is a tachistoscope and what it does is to flash up cards for a brief period of time. You will have time to read what is on the card but it won't appear for a long time. On each card there will be one word and what you have to do is to say whether it belongs to a specific category. The cards are arranged in groups of 5 (consecutively) and before each group I will say which category you have to look for. So your answer will be to say YES or NO; yes, it belongs to, say, the category of containers or no, it doesn't.

This thing here is a microphone which stops a timer as soon as you say yes or no, and what I am interested in is how long it takes you to decide whether the word belongs to that category. So, say as quickly as you can whether the word belongs to the category.

The procedure will be that I will say 'Right', and the word will appear in the next couple of seconds."

Any queries about the apparatus were dealt with and any procedural difficulties sorted out.

Three groups of 30 words were then presented with a break of 2 minutes between each group. The experimenter noted the RT.

In the final part of the session the category size estimator was given to find "how many words you know, as you cannot be expected to think up words if you have never seen them before."

Appendix 15

EXPERIMENT 6

	T-L Frequency	Category	Non-Category
Visual	AA	pay	feet
	AA	egg	manner
	AA	paper	nature
	A	pipe	judgement
	A	prayer	habit
	44	pig	observation
	26	jury	luxury
	25	guy	excess
	12	pang	violin
	8	gravy	bard
	4	jig	prow
	1	jag	drover
	X	pip	usurer
	X	yearning	cusp
	X	ping	hiccup
Acoustic	AA	touch	distance
	AA	teacher	door
	AA	march	word
	A	approach	aid
	A	porch	quality
	30	discharge	grief
	30	scratch	mate
	22	butcher	jacket
	11	wrench	pastime
	10	crutch	suffrage
	5	hunch	hemp
	1	escutcheon	cataclysm
	X	bitch	jinx
	X	sachet	corollary
	X	conch	arson
Semantic	AA	case	island
	AA	glass	youth
	AA	safe	example
	A	bottle	fault
	A	shell	harm
	36	tin	velvet
	26-50	file	observer
	15	cot	trumpet
	11	coffin	uproar
	8	hamper	parsley
	7	packet	severity
	3	casserole	oboe
	X	briefcase	cello
	X	haversack	eider
	X	pannier	manacle

Appendix 16

Experiment 6 - Decision time for presented category words grouped for Thorndike-Lorge frequency

VISUAL CATEGORY

Frequency		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
AA	Decision time (c. secs)	92	69	85	93	77	84	47	71	100	87	134	67	76	84	84
	Set Size	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	No. words generated	3	0	0	1	1	0	1	0	1	1	1	2	1		2
A	Decision time	93	68	72	108	60	82	50	64	88	91	110	73	100	67	65
	Set Size	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	No. words generated	0	0	1	0	1	0	1	0	0	0	0	0	0		1
26-50	Decision time	101	69	67	95	78	86	50	70	100	102	92	68	70	77	66
	Set Size	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
	No. words generated	0	0	2	2	1	0	1	0	0	0	1	1	0		2
11-25	Decision time	94	92	101	93	72	101	52	60	77	98	101	79	84	70	73
	Set Size	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
	No. words generated	1	0	0	2	0	1	1	0	1	0	1	0	0		1
1-10	Decision time	101	76	92	114	40	89	47	67	105	103	137	81	89	94	60
	Set Size	72	79	81	55	76	79	81	79	69	81	62	65	76	69	67
	No. words generated	2	1	1	1	3	2	1	0	3	1	0	0	0		1
X	Decision time	108	78	100	105	75	81	58	71	102	107	95	74	82	68	91
	Set Size	23	29	29	23	23	23	28	23	17	23	23	17	28	23	17
	No. words generated	0	3	1	0	0	0	0	0	0	0	0	0	0		0

ACOUSTIC CATEGORY		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
Frequency																
AA	Decision time (c.secs)	76	72	68	77	71	89	49	80	71	84	78	61	78	68	61
	Set Size	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
	No. words generated	3	2	1	2	1	2	0	1	0	2	1	4	1		2
A	Decision time	95	83	75	79	76	103	51	120	106	120	95	88	102	99	97
	Set Size	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
	No. words generated	0	2	1	1	1	1	0	0	1	0	0	2	1		1
26–50	Decision time	131	66	84	84	92	99	49	153	92	165	137	87	96	92	62
	Set Size	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
	No. words generated	2	2	4	1	3	0	0	3	4	1	2	0	4		2
11–25	Decision time	99	73	87	83	78	107	42	111	83	88	153	71	85	68	60
	Set Size	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
	No. words generated	4	1	0	3	1	1	2	1	1	2	1	2	4		1
1–10	Decision time	84	77	93	85	71	90	45	136	103	85	98	91	109	68	87
	Set Size	67	73	76	70	70	73	76	73	73	73	67	70	76	67	76
	No. words generated	2	5	1	5	5	3	5	2	7	4	3	1	1		7
X	Decision time	95	68	69	85	78	89	50	94	89	133	89	74	86	72	87
	Set Size	24	19	24	19	16	16	24	21	21	21	21	14	24	19	16
	No. words generated	1	1	0	1	2	0	1	0	1	2	1	0	0		0

Appendix 16

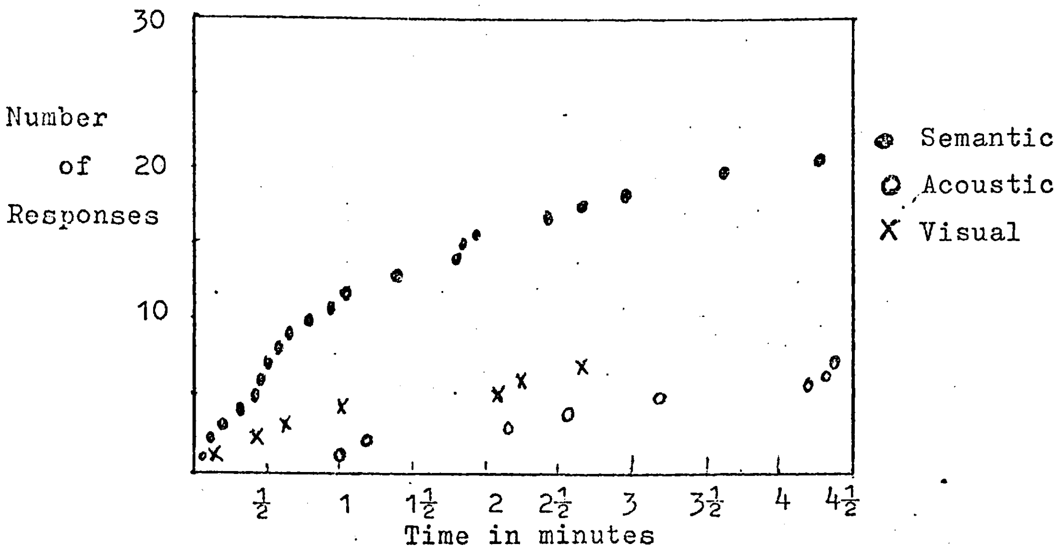
SEMANTIC CATEGORY		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
Frequency																
AA	Decision time (c. secs)	84	77	81	102	96	78	41	78	92	84	91	72	80	120	66
	Set Size	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	No. words generated	5	4	4	3	4	3	6	4	4	4	3	2	3		4
A	Decision time	64	80	102	87	120	70	59	68	81	81	114	86	106	69	91
	Set Size	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
	No. words generated	4	4	3	1	2	3	1	2	2	2	4	3	2		2
26–50	Decision time	76	78	77	76	93	81	49	71	78	85	101	72	77	81	60
	Set Size	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
	No. words generated	4	4	3	2	4	2	3	5	4	3	4	2	2		3
11–25	Decision time	124	78	89	90	125	83	77	90	78	97	101	82	82	74	56
	Set Size	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
	No. words generated	3	3	2	2	2	0	2	2	2	3	4	2	2		3
1–10	Decision time	83	64	80	91	77	74	64	70	101	95	81	106	80	78	70
	Set Size	65	67	70	54	63	61	70	67	67	65	61	67	65	63	65
	No. words generated	4	7	7	5	2	4	9	4	4	5	7	6	5		6
X	Decision time	90	76	76	83	135	80	55	84	88	86	88	105	68	104	68
	Set Size	18	23	28	11	25	20	11	18	16	18	18	18	28	16	13
	No. words generated	1	1	1	0	0	1	1	0	0	0	1	0	0		0

Appendix 18

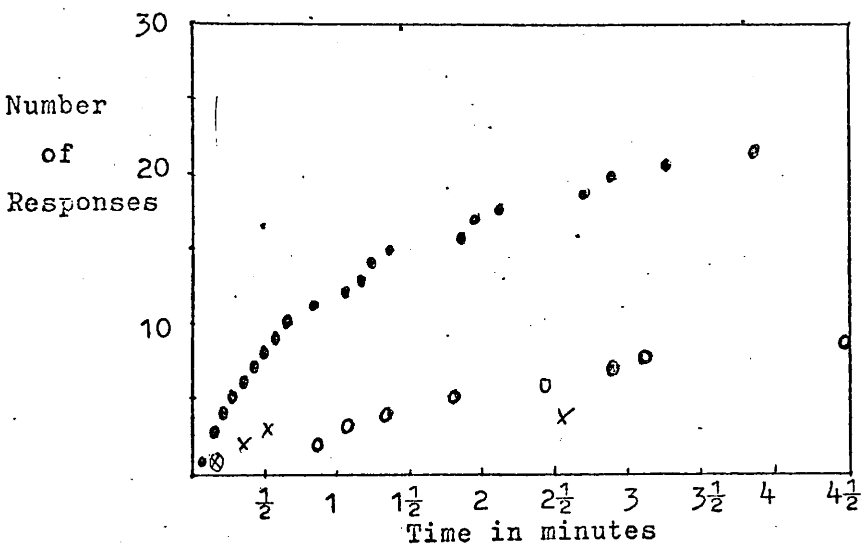
The findings presented here are provided as incidental to the main stream of the results in the rest of the report. Though they are of some relevance to the model proposed, it is acknowledged that the results are incompletely analysed. It had been hoped to illustrate that the derivation of the distance model could be made to fit the latency data, but on closer inspection the task seemed so immense that only the exploratory findings are presented.

A representative selection of graphs of the subject's output appear on the following pages. The minimum time interval considered was 3 seconds and a subject's accumulated output was plotted if he had produced a word in the preceding 3-second interval. Thus the subject's rate of output can be measured by the gradient drawn between any two adjacent points during the time course of the generation according to that code. Each of the three codes appears on each graph, and the final graph represents the total output of all the subjects together.

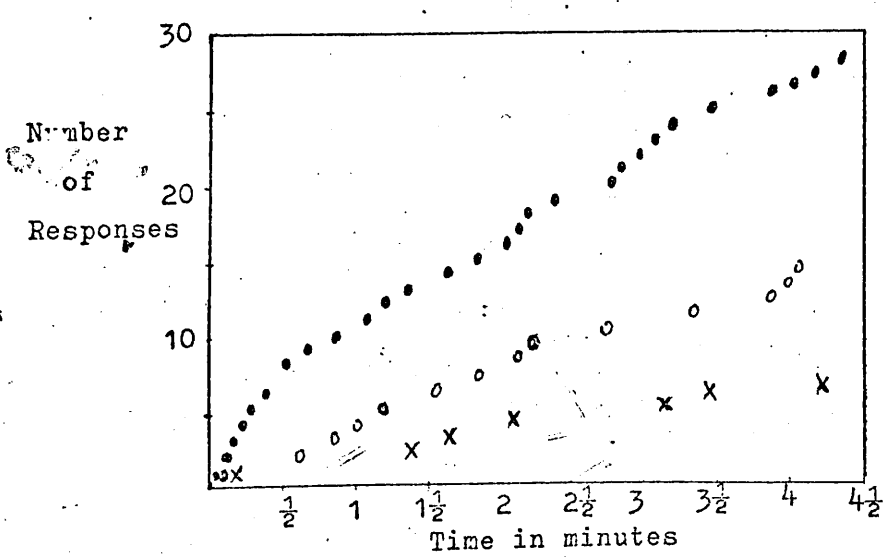
What is noticeable in all of these graphs of subject's output is that the gradient changes constantly and that these changes take place so as to radically increase the rate of output at various points. This means that the subject generates words very rapidly at the beginning of the time period and begins to slow down; after a varying period of time he discovers another suitable word and then



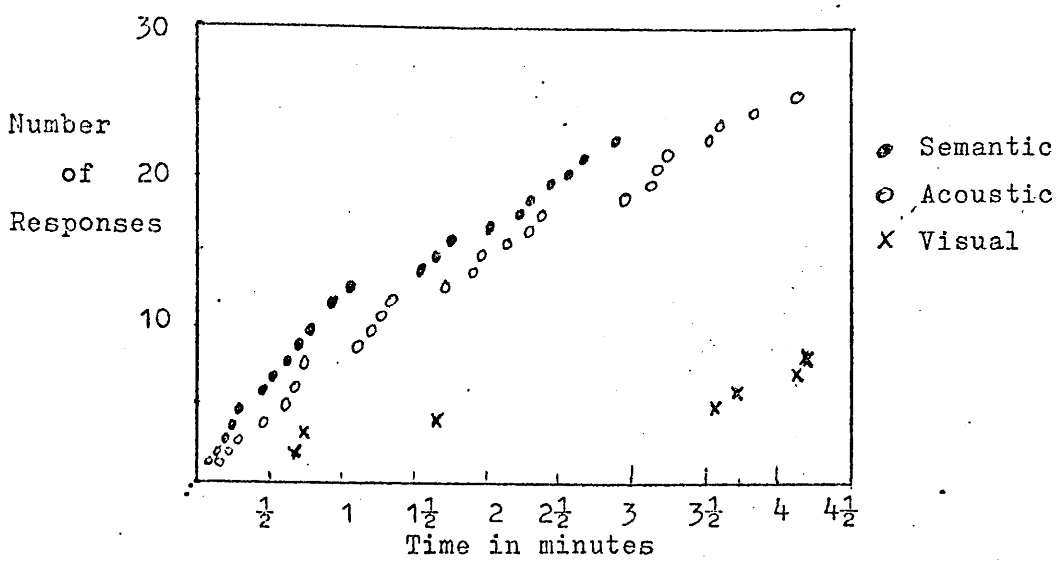
Subject No. 1



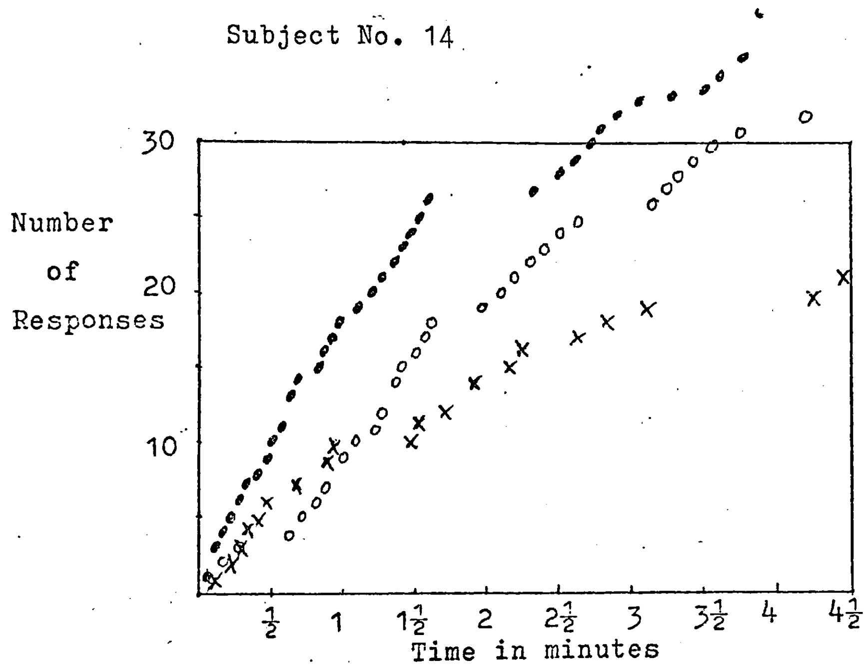
Subject No. 9



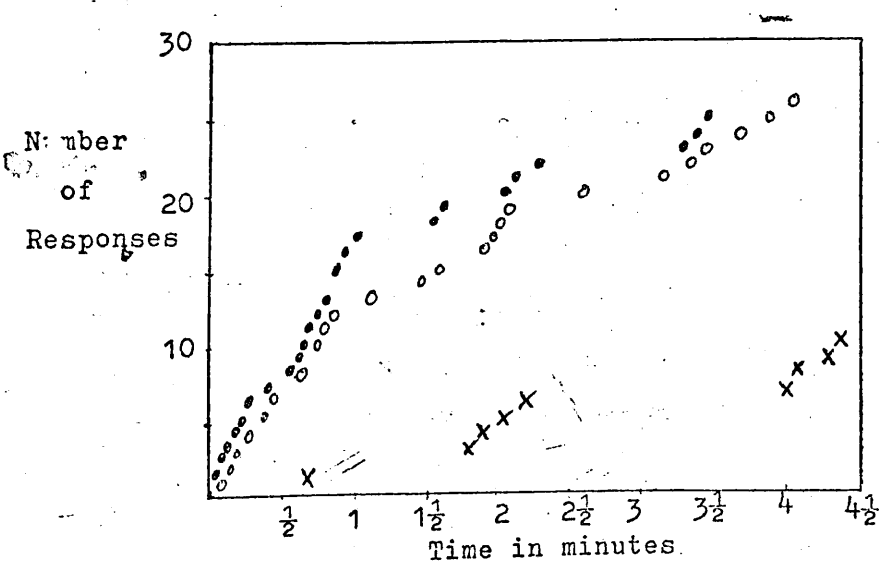
Subject No. 10



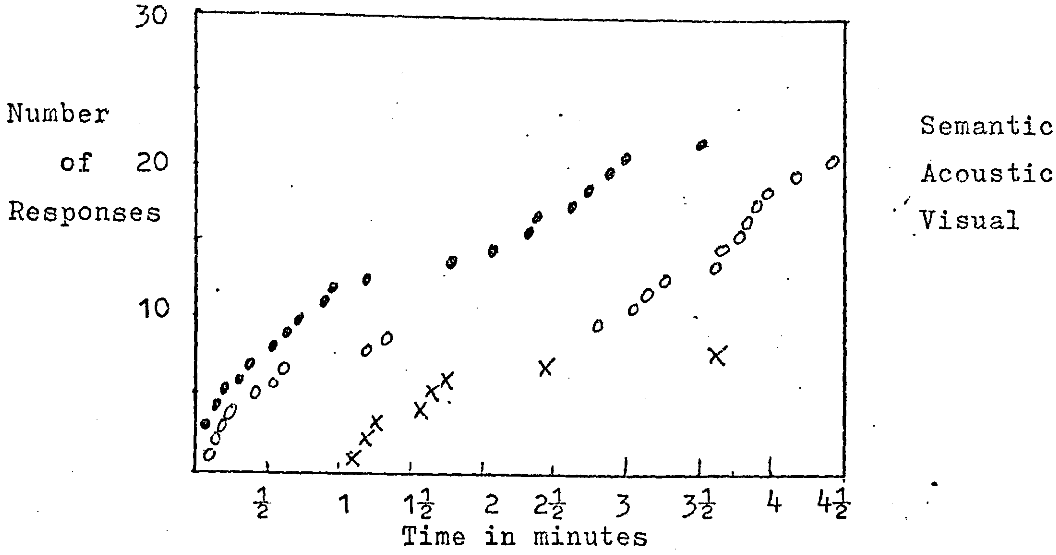
Subject No. 14



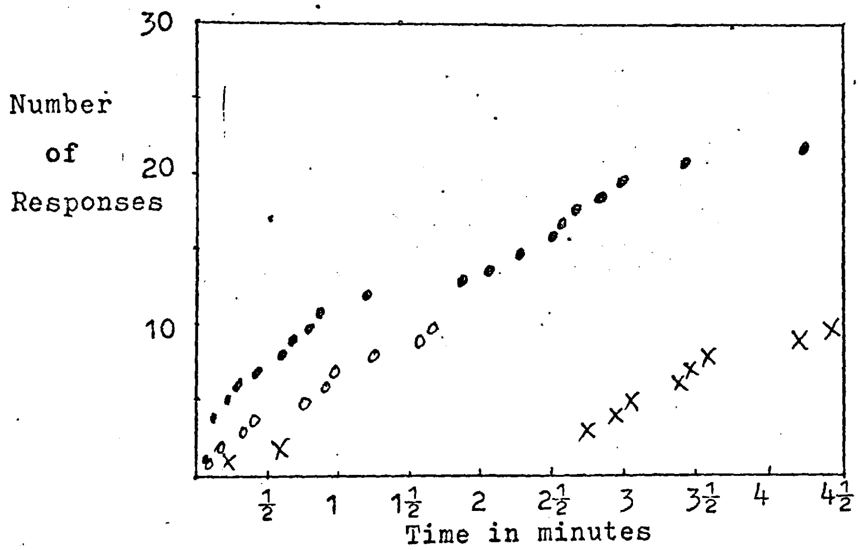
Subject No. 15



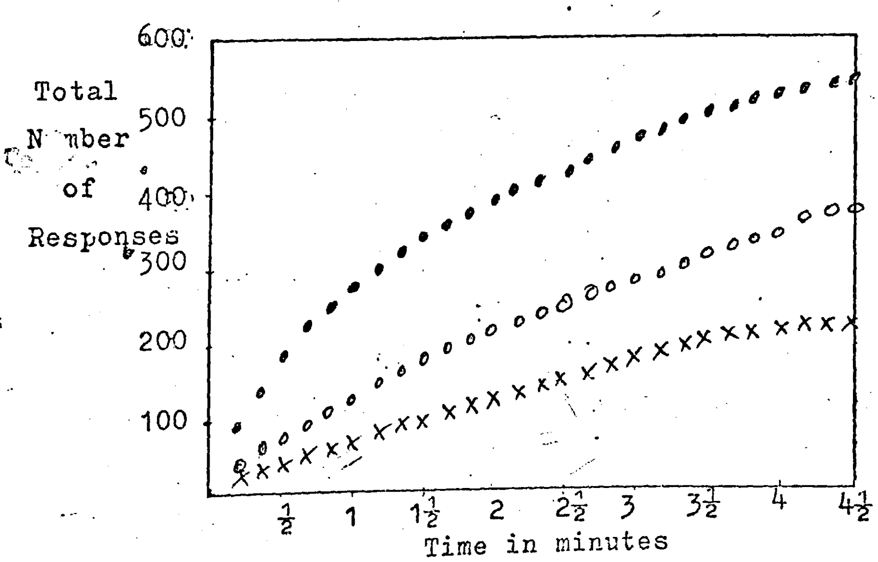
Subject No. 16



Subject No. 19



Subject No. 24



All Subjects - cumulative production.

suddenly produces other words at a hugely increased rate of output. He then slows down again (arrows indicate points on the graphs where this appears to occur). This variation weighs against a constant search process which would suggest a deceleration of output with time. Also it cannot be explained in terms of a lessening number of words which can be output, since no subject produced more than 30 words out of a total possible of around 180.

So why should there be such fluctuations in rate of output? The proposal that can be made on the basis of the model is that the rate of search fluctuates (though it may be number of searches in time, or the extent of the searches). If one could speculate a little, search might be characterised as a decaying exponential in rate of search or extent, such that when a word is found the rate increases momentarily thus increasing the chance of an immediate other output, which in turn would restore the search level. Working against this would be an almost additive inhibition so that the subject could not go on producing words from the same area of memory, at the fastest rate. As the words in a particular area are reported the search rate would naturally decrease to an asymptote (which will vary from person to person and may be a true reflection of his commitment to the experiment) from whence it will only be increased on the finding of a suitable word.

This type of process would allow the type of fluctuations seen in the graphs and particularly characterise the patterns for the more difficult categories like visual. An example of this is Subject 24, where he seems to suddenly find a "key" to words, having spent a long period without production. Subjects 19 and 16 are similar. Since the basic proposal of the system is the visual words are more widely spaced in memory (and it is supported by the lack of bursts of words produced together at the beginning of the time period) it is clearly not simply a matter of stumbling on a group of words which can be output quickly.

In many ways the whole idea is not counter-intuitive. Subjects often report a "mental block" whereby they are unable to retrieve words, no matter how hard they appear to try. It is also true from observation that one feels that the longer the subject goes without producing a word the less likelihood he has of finding one. Neither of these can be attributed to boredom or fatigue as the trial lasts only $4\frac{1}{2}$ minutes. In the model's terms the block is due to the decreasing search rate and to a possible inability to shift the search to a different area of memory which would be more profitable.

Some further information can be added to the topic from research reported. Broadbent (1973) in an investigation of the Miller (1956) hypothesis of 7 ± 2 items as optimal in memory or output, presents data on the production of words given a category name. His suggestion is simple:

the groupings of subjects in storing and recalling, and reproducing words tend to be in bundles of 3 rather 6 or 7. His idea is to propose a model of a working store which deals with units of 3 items, and in this respect, his purpose is different from that here, but he does report data where the pattern of output in terms of latencies has been mapped. The results conflict with Bousfield, Sedgewick and Cohen (1954)* who predict that average rate of response varies with the number of possible responses still remaining unemitted. The main point is that the largest run of short intervals (words produced quickly) is as likely to occur near the end as at the beginning of subject's output. He remarks:

> "These results mean that changes, with increasing time spent on the task, take the form of pauses of increasing duration between runs of responses made with normal speed." page 10

which is exactly the finding noted above, that groups of responses appear at different times during the subject's output.

Broadbent tries to fit this into a model of working storage based on the fact that items are encoded into small groups. While this can be satisfactory for presentation-recall situations, it is unconvincing for production from long term memory, where the above search model would account for the results better. The available clusters which Broadbent requires before output, would be so large

*Bousfield, Sedgewick and Cohen (1954), Am. J. Psychol., 67, 111-8.

in quantity that search for a particular suitable one (for the requirements of the task) once a relevant single word has been found, would needlessly complicate the procedure for retrieval. Subjects do produce words in isolation, so that it is not completely accurate to suggest that they attempt to produce only clusters. It seems more likely that the search process varies. The value 3 for groupings, may be a product of the parameter determining the build-up of inhibition or of the stopping rule which determines a change of search starting point.

Also the cluster hypothesis does not agree with his reported findings on the category of TV programs, which do not exhibit runs as the other categories did. Subjects found difficulty in producing the names of TV programs; one is tempted to compare this with the visual code in the above experiments, in that the cue given is not the most efficient one for the retrieval of the desired information.

Thus Broadbent's data while agreeing with that reported here, has been used to support different hypotheses but one feels that in the realms of long term memory a search model gives a better explanation.

Other work which has a bearing on the findings comes from 3 papers: Johnson, Johnson and Mark (1951), Kaplan and Carvellas (1969) and Kaplan, Carvellas and Metlay (1971).

They are linked by the fact that they all use the same mathematical model to describe word production. The model comes originally from Bousfield and Sedgewick (1944)* and consists of an exponential equation:

$$N = C(1-e^{-mt})$$

where N is the number of words written at time t, C is the upper limit that the curve approaches asymptotically, and m measures the rate of depletion of supply. The measure C, it is suggested, represents the total supply of the kind of words called for by the instructions. Bousfield and Sedgewick claim to have found some psychological justification for these parameters.

Some serious doubts arise about the treatment of the data in the 3 papers due to findings reported in this thesis.

The most serious difficulty arises from the development of the parameter, c or C, from a performance asymptote to the number of responses available to the subject. In the experiments reported in this thesis some attempt was made to measure the exact size of the set available to the subject for his production. This figure can be used to estimate the accuracy of fit of the model. Johnson et al. (1951) in fitting the data to the production of U.S. cities and animals, estimate the parameters from a curve drawn by eye. Kaplan and Carvellas (1969) estimate C from the asymptote reached by the group data after 15 or 25 minutes production time. Both reported some degree of success in the fit.

* Bousfield and Sedgewick (1944), J. Gen. Psychol., 30, 149-65.

The method used here was a minimum chi-square, since in some cases the subjects had not reached the asymptote that Kaplan's subjects had. The results are shown in the table below.

Fit of the Bousfield model to total group data.

Code	χ^2	C	p or m	Real c
Visual	8.77	367	.21	4144
Acoustic	16.32	479	.31	4200
Semantic	43.03	581	.62	4052

Real C is the total of each subject's estimate of the number of words from the set which he said he knew the meaning of. Clearly the fit is poor with the chi-square value highest for the semantic code, thus giving the poorest fit. One feels that despite the inadequacies of grouping the data in the above manner, in line with Kaplan and Carvellas, the parameter C does not reflect the number of words available to the subjects. The psychological basis of C is not upheld.

One further problem concerns a claim made by Johnson et al. that individual differences in the number of words produced is determined very largely by individual differences in the supply of words. This does not appear to hold according to the findings of Experiment 4 which shows that set size does not predict the extent of word production.

Kaplan and Carvellas (1969) follow a similar line generally, in trying to establish psychological support for the model of word production. They do however, make some modification in that they allow words that are not equiprobable. Their distribution takes the form:

$$p_r = p_1 r^{-s}$$

where p_1 and p_r are the probabilities of the first and rth ranking words, and s determines the rate at which probability decreases with rank. The difficulty is that the method used to test this was to estimate the p values from the probability densities gained from the latencies and then to rank these as r, in the equation. Following this the test was to plot log p against log r which should result in a straight line if the relation holds. Unfortunately, any set of random numbers when ranked and plotted against the rank on a log-log scale will produce almost a straight line with a gradient of around 1. So, this cannot serve as a test of the model unless rank can be estimated in some other way. While accepting the usefulness of the suggestion that words are not equiprobable it looks unlikely that the above equation offers much advance on the difficulties of the original.

Kaplan, Carvellas and Metlay (1971) study the effect of context on production and offer some advance, since they suggest that there are search processes occurring. However it suffers from the same criticisms as the Johnson et al. paper in the estimate of the subject's supply of words, and since the subject is hypothesised to search through this estimated number of words and his search rate is based on it, it is not likely to be completely realistic.

The result of all this analysis seems to be a misjudgement of the complexity of the situation and a lack of understanding of some other variable which determines output independent of the extent of words available. In the model proposed in this thesis the other variable appears as search modified by factors in the word itself, in the relative locations of the words and in factors of motivation and inhibition which determine the extent of the search. The Bousfield model is too simple to account for these factors.

Before closing this section, it may be useful to make some suggestion as to the testability of the proposed model. Generally speaking it does tie in with the data reported and it would be possible to add further information along the same lines. Ideally however, the model should be able to predict the pattern of output given a search area and information on distances and familiarity. One cannot pretend that we are at that stage of knowledge at the present, but nevertheless it might be worth considering how the model might unfold.

The type of information about the individual which would form the basis of word generation could be obtained in much the same way that Kiss acquires his association data. Given sufficient chaining by the individual relative distances between items could be built up. Familiarity information might be obtained from ratings by the subject. Neither of these would be completely satisfactory but would allow a limited test of the model. Given a starting point the model says

$$\text{Probability of a Response} = k\,\frac{\text{familiarity}}{\text{distance}} \text{ or } k\,\frac{\text{familiarity}}{\text{distance}^2}$$

After each search a series of probabilities would be thrown up and on the basis of these, a word is generated or not. If a word is generated the search will remain at the same level; if not, it will be decreased by an amount x. At the same time, no matter the outcome of the first search the probabilities of output would be decreased by y to represent inhibition caused by the search. y may be a constant, while the effects of x would follow the suggested negative exponential. A stopping rule may be adopted by the system such that after a certain number of searches or of unsuccessful searches, n, the starting point is switched to some other aspect of the task.

At any point in time the production of a specific item can be assigned a probability conditional on the items which have gone before. Presuming an individual works consistently, the result s at the beginning of a sequence can be used to generate the parameters to simulate the rest of the response pattern. Given the matrices to specify distances and measures of familiarity, and an assumption that the initial search time is $t = 1$

arbitrary unit, the parameters x and y can be estimated from the pattern of output and the value n will be linked to Broadbent's cluster size which is 3. Some idea of the model's working could be obtained.

In conclusion, the model exists to explain the data obtained, in a manner which might encourage consideration of aspects of the memory base and the process of search. It does have a theoretical base and it can generate testable hypothesis, though at the present time the qualitative aspects predominate.

Glossary

Continuous Association - repeated associations to the same stimulus over a fixed interval of time.

Familiarity - an internal indicator of the frequency with which a subject has used or encountered a word.

Fluency - the number of acceptable responses in a time interval.

Originality - the relative infrequency with which a specific item has been produced by the group. As a clearer relation emerged with Creativity, the first term, Unusualness, was replaced by Originality.

Originality of Words Generated - as above but with reference to words produced in response to specific cues.

Originality Index - a measure of Originality per item produced. It is calculated as follows

$$\frac{\text{Originality}}{\text{Fluency} \times 4} \times 100$$

After the re-analysis of Experiments 2 and 3 this became:

$$\frac{\text{Originality}}{\text{Fluency} \times 10} \times 100$$

Recall Originality - the Originality score of the words remembered from a previous session of word generation.

Unusualness - a measure of the infrequency of a response in terms of the total group's responses.

Unusualness Ratio - a ratio representing the Unusualness per item produced by the subject.

Unusualness Rating	- the rating by subjects or judges on a ten-point scale, of the infrequency of a word based on their own familiarity with it.
Words Generated	- the words produced, theoretically from long term memory, in response to specific instructions delimiting the category.