

LINKAGE IN EXTRA-SENSORY PERCEPTION

VOLUME I

by

M. C. MARSH.

(A thesis submitted to Rhodes
University for the degree of
Ph.D. in Psychology).

Rhodes University,
Grahamstown,
Union of South Africa,
1958.

CONTENTS.

Where the title of a section is followed by an asterisk a more detailed list of its contents will be found at the page reference given.

	<u>Page:</u>
List of Tables:.....	3
List of Graphs:.....	5
Abstract:.....	6
Acknowledgements:.....	8
Preliminary Remarks:.....	10
Introduction to the problems studied (*):.....	12
General & specific aims of the Rhodes Experiment:.....	25
How the data was obtained (*):.....	29
How the data was scored (*):.....	79
First Main Analysis: Differences in the frequencies of hits scored on Experimental and Control Series Originals (*):.....	95
Second Main Analysis: Examination of the data to find whether subjects tended to score significantly more hits on each Experimental Series Original during the week in which it was used as a target than in other weeks (*):.....	158
Third Main Analysis: The relationship of hit-scoring ability to personality factors (*):.....	168
Fourth Main Analysis: Fluctuations of the scoring rate after the introduction of Linkage Material (*):.....	268
Fifth Main Analysis: Qualitative evaluation of a sample of drawings returned by the subjects (*):.....	326
General conclusions drawn from an integration of all the findings:.....	382
Final reflections and suggestions to experimenters wishing to repeat the experiment.	389
Glossary of Terms and Abbreviations used in the report:.....	392
Bibliography:.....	398
Appendices (*):.....	404

LIST OF TABLES.

<u>Number:</u>	<u>Title:</u>	<u>Page:</u>
I	Showing the significance of the differences in Raw Score hits on Control & Experimental Series Originals for the combined scores of C & E.G.S. for various classes of hit. (C1-C20).	107
II	Summary of the significances of the differences in A hits scored on C & E.S.O. from various types of theoretically expected frequencies. (C21A-C24).	151
III	Summary of the material bearing on whether subjects tended to score more hits on originals during the weeks in which they were the targets than in other weeks. (Raw scores on E.S.O. only) (C25-C34).	165
IV	Ditto: Scores corrected for bunching: (C35-C44).	166
V	Correlations obtained between subjects' hit scores and their personality test scores. (C45-C54).	203
VI	Showing personality test score means for various subject groups classified in terms of uncorrected A hits scored. (C55-C72).	210
VII	Summary of the significance of the differences in personality test score means for various groups of subjects classified in terms of uncorrected A hits scored. (C55-C68).	214
VIII	Summary of the Bernreuter B3I Introversion Scores. Significance of the mean differences for various subject groups classified in terms of uncorrected A hits. (C65 to C72).	216
IX	Significance of the differences between mean Introversion scores of groups of subjects (a) suggesting titles in the Preliminary Frequency Experiment used in C & E.S.O. & (b) Scoring uncorrected A hits in the Main experiment. (C73-C76).	218
X	Assessment of the significance of the Bernreuter FIC and B3I personality score differences between (a) A subject group scoring 0 A hits & (b) A subject group scoring 2 or more A hits, using nonparametric methods. (C77-C80).	225
XI	Summary of the data from the Psychic Questionnaire. (Including subjects failing to return the full 50 drawings). (C88-C96).	261
XII	Ditto, but excluding subjects failing to return the full 50 drawings. (C97-103).	263
XIII	Comparison of the P values obtained in Tables XI & XII. (C88-C103).	265

Continued overleaf:

LIST OF TABLES

<u>Number:</u>	<u>Title:</u>	<u>Page:</u>
XIV	Linkage Effects: Significance of the Raw Score fluctuations for all weeks pooled. (C104-C131).....	302
XV	Linkage Effects: Significance of Raw Score fluctuations for weeks A,C,D & E pooled, with week B eliminated. (C132 to C159).....	305
XVI	Linkage Effects: Comparison of the significance of the Raw Score fluctuations with those of Scores corrected for Bunching, calculated on weeks A,C,D & E pooled, but eliminating week B. (On E.S.O. only). (C132-138, C146-C152, C160-C165 and C166-C171).....	306
XVII	Linkage Effects: Significance of the Raw Scoring Rate differences between week A and each other week separately. (C172-C199).....	312
XVIII	Linkage Effects: Comparison of significances obtained from Raw Scores and Scores Corrected for Bunching on E.S.O. only. Week A compared with each other week separately. (C172-C185 and C200-C211).....	313
XIX	Linkage Effects: Significance of the weekly scoring rate differences between C & E.G.S. (Raw hits on E.S.O. only). (C212-C218).....	321
XX	Classification of hits on Experimental Series Originals into idea and shape resemblances.....	376
XXI	Classification of hits on Control Series Originals into idea and shape resemblances.....	378

LIST OF GRAPHS.

<u>Number:</u>	<u>Title:</u>	<u>Page:</u>
1	Personality Material: Sheep-Goat Attitude Analysis: Showing the influence of the subjects' attitudes on the frequency of scoring A hits.....	233
2	Personality Material: Showing the A hit scoring rate per trial for three subject groups rated well-adjusted, averagely adjusted and poorly adjusted on the Washburne Subtotal Scores.....	244
3	Combination of Sheep-Goat ratings with Washburne Subtotal ratings: showing A hits scored per trial for various subject groups.....	248
4	Linkage Effects: Uncorrected A hits scored in each week of the experiment by E.G.S. & C.G.S. on E.S.O. & C.S.O.....	288
5	Ditto, for hit category A+B+C+D+E.....	289
6	Ditto, for hit category All F, all G + H3.....	289
7	Ditto, for hit category All F, G & H.....	289
8	Ditto, for hit category I + K.....	289
9	Ditto, for hit category I + J + K.....	289
10	Ditto, for hit category "All hits combined".....	289
11	Linkage Effects: A hits Corrected for Bunching scored in each week of the experiment by E.G.S. and C.G.S. on E.S.O. only.....	290
12	Ditto, for hit category A+B+C+D+E.....	291
13	Ditto, for hit category All F, all G + H3.....	291
14	Ditto, for hit category All F, G & H.....	291
15	Ditto, for hit category I + K.....	291
16	Ditto, for hit category I + J + K.....	291

ABSTRACT:

The aims of this investigation were firstly to find whether 371 subjects, mainly students at Rhodes University, would be able to reproduce target drawings made 470 miles away by an agent in Cape Town, by means of General Extra-Sensory Perception; secondly, to find whether any association existed between any such hit scoring ability shown and personality ratings derived from a battery of personality test material; and thirdly, to find whether providing the subjects with material designed to link them more closely with the agent would increase their scoring rate.

The 17,440 drawings returned by the subjects were randomised, and scored against a randomised set of 100 drawings consisting of 50 which had actually been used as targets, intermixed with 50 that were equal in difficulty but which had not been used as targets, and which were inserted merely as controls. Three independent judges assessed the subjects' drawings, and awarded hits in terms of title, shape and association. In the crucial title hits highly significant deviations in favour of the experimental target drawings were found, the control series of drawings showing no such effects.

It was found that the subjects' hits were distributed evenly throughout the whole 25 day period of the experiment for each target, and showed no tendency to occur more frequently in the week a particular drawing was being used as a target than in the other weeks when it was not.

When the subjects were divided into a high scoring and low-scoring group in terms of their E.S.P. ability. the high-scoring group proved to be significantly more extraverted (as measured by the Bernreuter B3I scale) than the low-scoring group, confirming a relationship

previously reported by Humphrey. In addition the data showed several other relationships, predicted by the work of previous experimenters, although these did not reach the .01 level of statistical significance.

To assess the effect of the Linkage Material the subjects were divided into an experimental and a control group. The Experimental Group was supplied with correct Linkage Material, the Control Group with incorrect Linkage Material, which they were led to believe was correct. The two sets of Linkage Material were equated in all other respects. The Experimental Group showed significant improvements in scoring rate when supplied with a photograph and description of the agent, and a probably significant improvement when supplied with squares of handkerchief which had previously belonged to him. When supplied with squares of colour, also appearing on the agent's target sheets, no significant increase in scoring rate took place on the target drawings, but a significant decrease occurred in hits scored on the control targets. By contrast no consistent significant fluctuations were shown by Control Group Subjects, leading to the conclusion that the correctness of the Linkage Material was a factor contributing to its effectiveness.

A qualitative examination of the hits scored by the subjects revealed that they tended to reproduce the concepts depicted by the target drawings, rather than the actual shapes drawn by the agent, suggesting that the hits were being produced by a telepathic process rather than a clairvoyant one.

The results of the experiment suggest that the subjects were assisted by the Linkage Material to **gain access** to the agent's mind as a whole, and not to particular items in it. This throws doubt on the correctness of Whately Carington's Association Theory of Telepathy, which the experiment was designed to test.

ACKNOWLEDGEMENTS.

This investigation would not have been possible without the co-operation, kindness and generosity of a large number of people. I would like to express my thanks to these.

The report in its present form owes much to the help and constant encouragement of Professor E.H. Wild. His acceptance of the topic as one worthy of empirical investigation made it possible to carry out the research in the Psychology Department at Rhodes University, and his constant support as the investigation progressed lightened my task in so many ways that it becomes difficult to find an adequate expression of my gratitude.

A band of helpers contributed to the smooth running of the experiment itself. Thus the Fleischer family: Len, Quez, Maureen and Yvonne spent long alphabetical nights analysing the Preliminary Frequency Experiment lists. Mr. Rorke acted as an efficient and conscientious Agent. Dr. W.H. Williams drew up the descriptive personality material that was issued to the Control Group Subjects, and Mr. J. Levin lent his photograph to this material to give it verisimilitude.

Messrs. J Benetar, H. Blagden, I.S. Dorrington, P.J. Frost, D.R.C. Marsh, M.Meyer, N.I. Robertson, G.B. Russell, G.J. Salvesen, J.W. Selby, and the Misses N. Cross, B.M. Croxford, M. Currell, E. Hayes, S. Molteno, M. Roose and J. Wright helped with the weekly collection and distribution of the test booklets to the subjects in their residences.

My mother, Mrs. L.F. Marsh, and my brother, Mr. D.R.C. Marsh spent months helping to correct the personality test material, and assess the drawings which the subjects produced. Here again, thanks are inadequate to do justice to the size of their contribution.

Financial assistance from the Parapsychology Laboratory at Duke University, U.S.A., and from Mr. A.E. Andersen, Mr. C.J. Hill, and a research grant from Rhodes University, enabled me to visit the Parapsychology Laboratory at Duke University, and to meet some of the prominent workers in the field of this research

in England and America, and thus benefit from their comments and criticisms. Among these were Drs. Greenwood, Soal, Thouless, Gardner Murphy, Alice Buck, Dingwall, Schmeidler, Mrs. Dale and Mrs. Goldney, and other members of the Societies for Psychical Research in England and America.

I would like to thank too the staff of the Parapsychology Laboratory at Duke, and others, for many kindnesses to me while I was with them. Special thanks are due to Dr. and Mrs. Rhine, Drs. Pratt, T.N.E. Greville, Betty Humphrey, Mrs. Esther Forster, Mr. Frazer Nicol and Mr. Malcolm Turner, both for assistance with the evaluation of the research itself, and for so generously taking me into their hearts and homes during my visit.

The task of producing this report was greatly lightened by the efforts of my wife, Audrey, who typed all the copies of the calculations, carried out many of the checking operations involved, and helped with the final duplicating. But for her good offices, and those of Miss N. Fowler, this report would have contained many more of my idiosyncrasies of grammar, spelling and punctuation than it does. With the aid of "Modern English Usage and "The Oxford English Dictionary" these two proof-readers gently but firmly drew my attention to errors of expression I was committing with equanimity. That not all of these have been eliminated is no reflection on their efficiency, for on occasion I have preferred to retain my own unconventional usage, feeling more at ease with it than with the usual form.

I am also grateful for much valuable encouragement, advice and assistance from colleagues at Rhodes University, and other friends who have helped in ways not here specifically listed.

During the course of this report I have drawn on many sources for quotations. I have taken care to indicate the source of each quotation as it occurs in the text, and would here like to thank these sources for their material.

Finally I would like to thank the 371 subjects, whose participation made possible this research.

PRELIMINARY REMARKS.

To assist the reader to find his way about this rather lengthy dissertation, it would be as well to start with a brief description of the routine methods of presentation which have been employed.

The contents page immediately following the title page gives the location of each main section of the report. In cases where an asterisk follows the title of the section, additional more detailed lists of the contents will be found at the page references given.

Each analysis ends with a summary of the conclusions drawn from it. In addition, a summary of all the material dealt with in the dissertation will be found at page 382, and a short abstract on page 3 will already have been encountered.

For ease of reference each data sheet bears a serial number prefixed by a D, and each calculation is given a serial number prefixed with a C. Only one example of each type of calculation is given in the body of the report, as an illustration of the method used. The remainder of the calculations in any particular series will be found in the appendices at the end of the report. Much of the basic source data used for the calculations will also be found in the appendices, so that any reader wanting to check the calculations will have all the necessary material at his disposal. The bulk of the mass of raw material gathered precludes its inclusion in toto, but wherever possible it has been included in summary form. Supplementary data sheets will be found from time to time in the body of the report as well, as convenience dictates.

In some of the calculations it has been necessary to round the figures off to a required number of decimal places.

Some ambiguity exists among statisticians as to what rule to adopt when the next significant figure is a 5. In these assessments the procedure has been used of discarding the 5 if by so doing the rounded number would end with an odd number. The 5 is counted as the next highest unit if this would result in the rounded number ending with an odd figure. An example will make this clearer. Suppose we have to round off the following correctly to two decimal places each. Thus 5.375 would be written 5.37, but 5.385 would be 5.39. A few exceptions to this rule have been allowed in calculating the theoretically expected frequencies in calculations of the Fourth Linkage Analysis.

Each calculation has been done twice, the second time as a check, and all transcription has also been checked.

The findings resulting from the calculations are presented in Tables in the body of the report.

INTRODUCTION.

(1) The development of Psychical Research.....	13
(2) The background to the present research.....	16
(a) The contribution of Dr. Hettinger.....	16
(b) The contribution of Mr. Whately Carington.....	20
(c) The contribution of Dr. Betty Humphrey & other workers.....	24

INTRODUCTION.(1) The development of Psychical Research.

One of the most insistent needs of human beings down the ages has been the need to communicate with others. When people are separated by distance this need is difficult to satisfy, and a problem is created that man has attempted to solve in two main ways. The first attempt at a solution has been through inventions of a physical nature, such as writing, the invention of the telephone, radio and television. On such techniques is based our whole modern and effective system of communications. The second type of approach to the problem has been based on the belief that man has within himself the ability to communicate with others at a distance in ways other than through the normal senses. It is thought that the use of this ability depends on certain laws and techniques, which can be discovered and used. Throughout history the greater part of mankind has accepted and acted on this assumption uncritically, and a vast structure of magical techniques has been developed which is designed to put such powers of extra-sensory perception to effective use.

Western civilization has, however, tended to discount the effectiveness of such techniques, without putting any of them to experimental test. Thus many scientists would dismiss the possibility of E.S.P. existing, not on any scientific grounds, but because the topic arouses in them strong conscious and unconscious resistances. Tyrrell (93), and especially Ullman (94) among others have analysed the nature of these resistances. It is thought that they may spring in part from the feeling that Science owes its modern achievements to the fact that it shed the magical preoccupations of the Alchemists, and turned its attention instead towards physical causation. It is understandable that scientists holding this view can easily see in the investigation of E.S.P. a tendency for Science to return to the old miasmatic unproductivity of

Alchemy, and resist it. While it is possible to sympathise with such a feeling, the writer personally feels that such a danger is over-rated. An important advance in Science has been the development of valid and refined experimental methods. If use is made of such methods, the investigation of magical concepts is no threat to Science, since if they are false the experimental method will reveal them to be so, and there will be nothing further to investigate. If, however, research does demonstrate that men have such abilities as E.S.P., the importance of such a finding would be very great.

Some scientists have felt that E.S.P. is a legitimate field for serious research, and have undertaken investigations that have already yielded a harvest of findings of such bulk, and in so many languages, that a summary here is impossible. It is intended, therefore, merely to indicate the trend research has shown, and then deal in more detail with a few specific investigations that have been the inspiration for this research.

As with most branches of scientific knowledge, interest was attracted first by spontaneous cases which suggested that some people did have E.S.P. ability. In 1882 the Society for Psychical Research was founded in England for the purpose, among others, of collecting and investigating such cases. Similar societies were founded in other countries. At first the approach tended to be legalistic: investigators were concerned to establish that the event in question had actually taken place as described. Thus much attention was given to the reliability of witnesses, to the neglect of factors shedding light on the modus operandi of the occurrence.

Many of the early members of the S.P.R. were interested in mediumistic phenomena, and a large part of the early "Proceedings" are occupied by reports of such cases. Reports of controlled laboratory experiments are correspondingly rare. Then too, few of the early workers were able to give their uninterrupted time to psychical research.

Since the early descriptive beginnings, the emphasis has shifted more and more to controlled laboratory experimentation, and the statistical assessment of results. A big impetus to the change was given by the establishment at Duke University of a Parapsychology Laboratory. In the course of time this laboratory has acquired a full-time research staff, able to devote their whole energies to psychical research. Under the directorship of Dr. Rhine, the laboratory has been able to develop experimental material and apparatus, techniques of experimentation, and statistical methods of assessment that are very widely used by parapsychologists today. In addition, the experimental work carried on by the Parapsychology Laboratory resulted in a great increase in the number of precisely established facts concerning E.S.P. and P.K.

At first researchers were concerned to establish whether some individuals possessed the ability to obtain knowledge other than through the normal sense channels. The early research findings, which suggested that this was in fact the case, aroused a great deal of interest and criticism from scientists in other fields. Part of this criticism was constructive and informed, and was of great value, leading to improvements in experimental methods, making possible such carefully controlled investigations as the Pearce-Pratt series (57), which established the crucial fact that E.S.P. did actually exist. Researchers then turned their attention to a study of the nature of E.S.P. and the features affecting its occurrence.

With the improvements which were introduced in the experimental methods came increasing acceptance of the findings by scientists outside the actual field. This is reflected in the surveys undertaken by Warner (96) of the attitudes of American Psychologists to E.S.P. He finds that although the majority of the American Psychologists sampled are unconvinced that the existence of E.S.P. has been demonstrated,

the numbers of those holding open-minded attitudes are increasing. Universities have also accorded increased recognition to psychical research, and recently a Chair of Parapsychology has been established at Utrecht University in Holland.

(2) Background to the Present Research:

The investigation reported in this thesis was planned by the writer while serving with the Aptitude Tests Section of the South African Air Force from 1943 to 1946. It was carried out at Rhodes University, South Africa, during 1947, and the data analysed during the following years in the writer's spare time.

One of the first decisions that had to be taken concerned the kind of target material that should be used in the experiment. It was decided to use free drawings for this purpose, rather than the M.S.P. cards developed by Dr. Rhine and his associates. It was felt that though drawings would be much more difficult to assess statistically than would experiments using M.S.P. cards, a study of the data from drawings would shed more light on the actual telepathic process than would data from cards. In taking this decision the writer was influenced by the work of Hettinger and Whately Carington. Since the present research is an extension of the work of these two investigators, it is necessary to start by summarising their methods and findings.

(a) The work of Dr. Hettinger:

Dr. Hettinger has reported two series of experiments which he carried out at Kings College, London. From the first of these experiments Hettinger concluded: "The existence of an ultra-perceptive faculty may be ranked, at least, as a probability" (25) and: "When a Sensitive concentrates on an object belonging to a Subject a mental rapport is established between the Sensitive and the Subject, and whatever is perceived by the

Sensitive in virtue of the ultra-perceptive faculty is present in, or is accessible to, the psychological makeup of the subject at the time of the test". (25).

The method adopted by Hettinger in his second series of experiments was as follows. He stationed a transmitting agent and a receiving sensitive several miles apart in London. The task of the sensitive was to try and get telepathic impressions of the contents of the agent's mind. The sensitive was kept in ignorance of the whereabouts of the agent whose mind she was trying to contact, and of his identity, but was given personal objects belonging to the agent in the hope that they would help her to make this contact. Agent and sensitive synchronized watches from radio time signals, and at an agreed time the agent began glancing through an illustrated periodical previously chosen at random from a bookstall on the way to the experimental room. As he did this he noted down the time at half-minute intervals on the margins of the pages opposite the material he was cognizing. Simultaneously the sensitive began dictating to Hettinger any impressions that were coming into her mind, the exact time she made each statement being noted.

Hettinger then compared the statements with the material being viewed by the agent at the time the statements were made, to discover whether there were any correspondences between the statements made and the target material. He found some very striking resemblances. For example, at the exact moment when the agent was looking at an advertisement showing shoes with very prominent toes, the sensitive said: "Someone has peculiar shoes, something to do with the toes." (25).

To find whether his correspondences could be accounted for on the basis of chance or coincidence, Hettinger attempted a control matching of target drawings against statements made on other occasions when they were not being used as targets. He claimed that the resemblances so obtained were much fainter and fewer than those obtained in the actual experiment, when

statements were matched against statements at which they were actually aimed. He concluded that a genuine ultra-perceptive faculty was responsible for his results.

From a study of his results, Mettinger concluded that besides indisputably correct correspondences, other types of resemblances also occurred, and he classified these remoter resemblances into the following types:

- (i) Correspondences with slight additions or distortions.
- (ii) Cases where emotions, feelings or sensations were perceived by the sensitive.
- (iii) Movements perceived.
- (iv) Only one constituent of the picture noted.
- (v) A number of specific constituents seen.
- (vi) Cases of specialization and generalization.
- (vii) Illusions, where the sensitive misinterpreted the perception.
- (viii) Perception of an association of the picture.
- (ix) Relationships shown in the picture noted.
- (x) Cases of special interest, cases concerning reading matter, and "far-fetched" cases.

His results suggested that any experimental procedure should not merely assess clear correspondences, but should allow for the possibility that the telepathic message might be modified and distorted before reaching consciousness.

Warcollier, a French worker, had made such modifications the subject of special study, and an English

translation of his work (95) in 1947 shed much additional light on the nature of such distortions and changes.

Hettinger's work has recently been severely criticised by Scott (71) on the grounds (among others) that

- (i) Hettinger tends to generalize from isolated examples.
- (ii) There is a possibility that his selection of correspondences, which seem so impressive, might still only be due to chance.
- (iii) His control experiment is not watertight: Scott argues that since Hettinger knew which remarks were aimed at which targets, he might have been biased in his assessments.
- (iv) That his statistical procedures were not watertight.

In rebuttal Dr. Hettinger points out that the nature of his results are such that by inspection they are convincing, and do not need statistical proof of their validity. Not only do the sensitive's remarks correspond to the pictures at which the agent was looking, but also they were made at the exact time the agent was looking at the picture to which they referred, and thirdly, the sequence of remarks as a whole corresponded to the sequence of the target material. (71).

In his book (25) Hettinger analysed the types of linkage that took place between his agent and sensitives, and concluded that a linkage object was an essential part of the telepathic process. Although his work is not above criticism, it suggested to the present writer:

- (i) The value of using free drawings as target material.
- (ii) The need for evolving a flexible scoring scheme that would take into account

possible distortions of the telepathic message.

(iii) The need to establish the exact function of the object that was psychometrised.

(iv) The need to find a watertight statistical method of assessing the results of such experiments.

(b) The Contribution of Whately Carington:

During 1940 Whately Carington published a report of a most painstaking and careful research using drawings as test material. He used a statistical method that had met with the approval of Dr J.O. Irwin, Mr. Oliver Gatty, and Prof. C.O. Broad. Carington claimed that he had at last succeeded in evolving a repeatable experiment, and suggested that other experimenters test this by carrying out similar experiments. The dislocation of scientific research and communication caused by the war caused some delay before his suggestion was acted upon, but in 1945 Taves, Murphy and Dale published an account of American experiments on the paranormal cognition of drawings which tended to confirm Carington's findings. (90). In 1947 West published an article on "Mass experiments on the Psi-Cognition of Drawings"(98). During the following year Schmeidler and Allison reported a repetition of Carington's experiments (66). These publications were unfortunately not available to the writer at the time the Rhodes Experiment was carried out, but a careful search of the publications available failed to reveal any serious criticism of Carington's method, which will now be briefly described.

In his experiment, Carington and his wife acted as the agents, and made target drawings in Cambridge, England. Some of the subjects were as far away as Duke University, in America, but the majority were living in England. These subjects tried to reproduce the target drawings by means of E.S.P. The drawings made by the subjects were afterwards

compared with the originals, and their successes compared with the number which would be expected from chance alone, to see if the difference reached statistical significance. The following summary is taken from Carington's report (15):

"Five experiments in the 'paranormal' cognition of distant drawings have been conducted. In each of these, ten different originals were used, and ten fresh originals were used in each experiment. All originals were selected by a substantially random method. About 250 percipients of both sexes, producing about 2,200 drawings, took part. Very few percipients, probably not more than ten or twelve, took part in more than one experiment. In no experiment was there the smallest possibility, humanly speaking, of any percipient obtaining any clue to the nature of any original by normal sensory means or by rational inference. The drawings were marked against the whole fifty originals by a judge who had no clue or information as to which originals were used in which experiment. A total of 1,209 drawings were found which were judged to be sufficiently like one or other of the originals to deserve mention. From the data it is possible to calculate how many of these resemblances or 'hits' would be 'winners', i.e. hits on originals used in their own experiment, if chance alone were operative, and how often this value would be exceeded by any given amount. It is found that the excess is such as would be equalled or surpassed only about once in some thirty thousand such investigations if chance alone were responsible. In other words, percipients' drawings resemble the originals (considered as a group) at which they are aiming more closely than they resemble originals at which they were not aiming to an extent which cannot plausibly be attributed to chance."

"Examination of the data from another point of view shows that these resemblances do not occur exclusively, or even most often, at the same time as the display of the original concerned. But there is a fairly regular tendency for drawings

which resemble a given original to occur relatively more frequently on occasions which are near to that on which it was displayed than on others which are more remote. This effect is observable to a significant extent in both directions."

"The main conclusions indicated by the facts are, first that there is a real cognitive relation of some kind, (direct or indirect) between percipients and originals, second that this may be either of precognitive or retrocognitive form."

It should be noted that while Carington's results were highly significant statistically, the total effect due to E.S.P. was very small, since 2000 drawings yielded only 50 drawings above chance expectation, or roughly a 2½% effect.

Carington used no psychometric objects similar in nature to those used by Hettinger, but instead gave the majority of his subjects a photograph showing the relevant parts of his study where the target **originals** were exposed. Carington says that this photograph was intended merely as a placebo: he did not expect it to act otherwise than by suggestion, so did not specifically design his experiment to investigate the effect due to it.

Later Carington published a book (14) in which he put forward a theory of how telepathy takes place. This theory has become known as the "Association Theory of Telepathy," and has been quite extensively criticised on theoretical grounds, unlike his experimental method, which has not been called in question. Since part of this investigation is designed to test Carington's Association Theory, it needs to be summarised very briefly here. It will be dealt with in more detail later, in the "Fourth Main Analysis".

Carington suggested that telepathy was essentially associative in nature, and depended on common ideas present in the minds of both agent and subjects. He termed such a common idea a "K Idea". These K ideas formed the essential link between agent and subjects. He postulated that any

association formed in the agent's mind with such a K idea, was automatically formed in the subjects' minds as well.

In his experiment Carington suggested that the "Idea-of-the-experiment" was the K idea present in the minds of agent and subjects. When making the target drawings the agent tended to associate them with the "Idea-of-the-experiment". The essential assumption of the theory is that this association is formed at the same time in the minds of the subjects as well. When they sit down to make their drawings, this "Idea-of-the-experiment" comes uppermost in their minds, and brings with it the drawing that the agent has associated with it.

Carington produced evidence to show that his results conformed to the laws of Recency and Frequency in associative phenomena. The theory is fully set out in chapter VI of his book. (14)

It seemed to the writer that it would be possible to test Carington's theory of K ideas, by finding out whether the results obtained from a telepathy experiment conformed to the law of Vividness of presentation in the formation of associations. On the basis of this law, one would expect that the more vivid and specific the K ideas provided, the better should be the subjects' performance. Since Psychometry can be regarded merely as the provision of a concrete K idea, it would seem appropriate to test both the value of K ideas, and of psychometric articles at the same time.

In order to check Carington's claim to have evolved a repeatable experiment, it was resolved to follow his procedure as closely as possible. Consequently the instructions which were issued to the Rhodes subjects follow those of Carington very closely. The Rhodes experiment was planned so that the same type of statistical assessment of the data could be made. However instead of having five separate experiments, each using different subjects, the Rhodes experiment was divided into five sections, but

the same subjects were used for all five sections. This difference is important, and should be borne in mind, as it appears to have had repercussions on the Rhodes results.

It will be remembered that Carington found that his results were displaced precognitively and postcognitively in time. It seemed to the writer therefore necessary to adopt a method of assessment in the Rhodes investigation that would allow for such displacements in time. Though the procedure finally adopted in this investigation differs in some respects from Carington's, notably in the way the originals were selected, yet the format and inspiration comes largely from Carington, and the writer is very conscious of his indebtedness to Carington. At the same time it was felt that certain improvements could be made to Carington's technique, and these will be explained as they arise in the course of this exposition.

Experience in the Aptitude Tests Section during the war had impressed the writer with the need for using large groups of subjects if reliable results were to be obtained. It was therefore resolved to use as large a group of subjects as possible. Thus the subjects in the Rhodes experiment produced 17,440 drawings, as against the 2,200 produced by Carington's subjects.

(c) The contribution of workers in the Personality Field, such as Dr. Betty Humphrey:

It also seemed useful to carry further the work started by previous investigators such as Humphrey, Stuart, and Schmeidler, in attempting to correlate E.S.P. scoring ability with specific measures of personality. It was therefore decided to administer a battery of personality tests to the subjects before the actual experiment started.

Having now surveyed the chief sources of inspiration for the Rhodes experiment, we can proceed to a summary of its general and specific aims.

THE GENERAL AND SPECIFIC AIMS
OF THE RHODES EXPERIMENT.

- (1) Definition of the problem:.....25
- (2) The general aims of the Rhodes Investigation:.....26
- (3) Specific aims of the Rhodes Investigation:.....27

(1) Definition of the Problem:

A general survey of the literature of psychical research, and the findings surveyed in the preceding chapter, convinced the writer that the existence of E.S.P. had been successfully established. But the work reported raised many new problems. One of the most interesting of these seemed to be the nature of the relationship between the agent and the subject which made possible the occurrence of psi action between them. It would be of great value to know how the subject is able to select from all the possible material available to him just those aspects that relate to a particular agent and a particular target.

Psychical researchers have used certain methods, such as psychometry, in the belief that they help to establish links between people which facilitate paranormal action between them. However, practically no empirical evidence is available as to whether such procedures are actually effective in putting the subject in contact with the agent or not.

This problem, dealing with the nature of the relationship established between two people in a psi process, the writer has called the Linkage Problem.

(2) The general aims of the Rhodes Investigation:

Though some of the research findings, such as the effect of emotional rapport (42, 49) and genetic similarity (54) bear on the linkage problem, yet there seems to have been no attempt made to establish whether the E.S.P. scores of a large unselected group of subjects are better when they are given linkage material than when they are not. An investigation of this problem would appear to be fundamental to the whole problem of linkage.

It was therefore decided, first, to try and establish whether some of the commonly accepted linkage techniques had

any value or not. If some of these techniques proved to be effective, it was planned secondly to attempt to establish the relative effectiveness of the different types of linkage. Thirdly, it was decided to try and isolate some of the qualities which contributed to the effectiveness of any of the techniques which proved promising. Since such an analysis can only be valid on data genuinely produced by the subjects' E.S.P. ability, it was decided to adopt a procedure that gave fair promise of producing such data. Carington's method was the one chosen as the one most likely to fulfil this requirement. It was resolved that the experimental results would be analysed first to find whether or not they showed any effects attributable to the operation of E.S.P.

It was decided to test three types of linkage techniques: psychometric links, verbal and symbolic links, and K idea links such as postulated by Carington's theory.

Another aspect of the linkage problem is that dealing with the personality traits shown by good and bad E.S.P. subjects. If more were known of the kind of person who finds it easy to establish psi contact, considerable light would be thrown on the nature of the linkage itself. Consequently it was resolved to administer a battery of personality tests to the subjects taking part in the experiment and to correlate any E.S.P. success shown with the personality test scores.

(3) Specific aims of the Rhodes Experiment:

(a) To find whether a group of subjects in Grahamstown, South Africa, could gain paranormal knowledge of drawings made by an agent in Kenilworth, Cape town, South Africa. The Trigonometrical Survey Section, Bloemfontein has kindly calculated that these two ~~centres~~ lie approximately 470 miles apart, as the crow flies.

(b) To find whether any such E.S.P. ability shown

was correlated with the personality test scores of the subjects.

(c) To find whether linking the subjects to the agent in various ways would affect their rate of scoring.

HOW THE DATA WAS OBTAINED:

(1) Introduction.....	30
(2) The subjects used.....	30
(3) The battery of personality tests administered.....	30
(4) Construction and administration of the Psychic Questionnaire.....	31
(5) The Preliminary Frequency Experiment.....	32
(6) The selection of Control and Experimental Series Originals.....	36
(7) The time sequence of the experimental sections.....	37
(8) Procedure followed by the agent.....	37
(9) Instructions issued to the agent.....	37
(10) Procedure followed by the subjects and experimenter.....	43
(11) General instructions issued to the subjects.....	44
(12) The use of Control and Experimental Groups of subjects to investigate the influence of the Linkage factors introduced.....	49
(13) The linkage factors provided in	
(a) Section "A".....	51
(b) Section "B".....	51
(c) Section "C".....	54
(d) Section "D".....	54
(e) Section "E".....	63
(14) The specific instructions issued to the subjects on how these linkage factors were to be used.....	69
(15) Diaries kept by the subjects, agent and experimenter.....	76

(1) Introduction:

In this section it is proposed to undertake a concise survey of the experimental procedure carried out. As far as possible the actual instruction sheets issued to the various participants will be included, but in some cases it has been necessary to retype the material, so as to leave a binding margin. Consequently some forms may appear slightly more cramped than they were in reality when issued.

(2) The subjects used:

Before appealing for subjects, interest was enlisted by the writer giving lectures on E.S.P., interviewing possible subjects personally, and creating a feeling among the subjects that they were being given the opportunity to take part in something in which they were personally interested. As a matter of policy this personal contact was maintained with as many of the subjects as was possible throughout the experiment. By these means 371 subjects were collected who were willing to take part in the experiment. With only a few exceptions these were students in residence at Rhodes University College, Grahamstown, South Africa. The exceptions were a few members of the university staff, and a few interested outsiders who asked to take part. All except three of the subjects were resident in Grahamstown. Of those not living in Grahamstown, one was a medium living in Johannesburg, and the other two were on holiday and in transit during considerable parts of the experiment.

The age range of the group was from 16 to 65 years, inclusive, with the great majority of the subjects in the 17 to 21 year old classes. (A table showing the age distribution of the subjects can be found in the appendices).

(3) The administration of the battery of Personality Tests:

Before the main experiment started, a battery of personality

tests was administered to the subjects. The tests chosen were:

The Washburne Social Adjustment Inventory, Thaspic Edition. (97)

Pressey's X-O Tests. (50)

Allport Vernon Study of Values. (3)

Bernreuter's Personality Inventory. (7)

Of these, the Washburne was included as a measure of rapport with the environment, and over-all psychological adjustment. The Pressey was included in the hope that it would provide a measure of the fluidity and richness of the subjects' associations, as well as measures of certain psychotic and neurotic tendencies. The Study of Values was included in order to get measures of the subjects' broad attitudes to life, which it was hoped might be found related to scoring ability. Finally the Bernreuter was included because Humphrey had reported that some of its measures successfully discriminated between high-scoring and low-scoring groups of subjects. (37) These tests will be more fully described in the "Third Main Analysis" that follows later.

(4) The construction and administration of the Psychic Questionnaire:

A study of the material available on psychical research suggested that psychical ability might be related to a large number of miscellaneous factors, such as history of psychic experiences in the subject or his immediate relations, the predominant type of imagery used by the subject, belief of the subject in psychical phenomena, etc. In addition, it was thought that it would be interesting to sample opinion in a university group of students on psychical phenomena in general, and to discover to what extent the various types of phenomena were accepted as genuine. A Psychic Questionnaire was therefore constructed. Among the questions included was that asked in the Census of Hallucinations reported in the S.P.R. Proceedings. (74)

This was included in order to find out how common hallucinations of various sorts were in the university group, and whether these were related to any E.S.P. ability that might be shown.

This too will be dealt with fully in the "Third Main Analysis".

(5) The Preliminary Frequency Experiment:

One of the first questions that had to be decided was what target drawings were to be used in the experiment, and how they should be chosen.

In his experiment Whately Carington selected his originals by opening a dictionary at random, and making a drawing of the first object on that page that could be drawn. It seemed to the writer that this method was open to a number of objections. The most serious of these was that no control was possible over the popularity of the originals so chosen, so that Carington might have a large chance factor entering in to dilute his results. For instance, if the choice of an original happened to fall on "car", large numbers of the subjects might be expected to draw cars from chance alone, as it is a very common object, and this chance scoring might tend to obscure any real E.S.P. effect that the data might contain.

Carington also compiled a catalogue of the frequencies with which the various drawings had occurred in his experiments, & suggested that subsequent experimenters could use such a catalogue to establish the chance frequency with which any object could be expected to occur. This procedure too, seemed to the writer to be open to many objections. It seemed quite unjustifiable to assume that subjects drawn from one culture would tend to think of objects with the same frequencies as subjects from a different culture. Seasonal and chance factors might also affect the frequencies.

Instead it was decided that more valid frequencies would be obtained if the mental content of the actual subjects

who were going to take part in the experiment was sampled just prior to the start of the actual experiment. Thus a few weeks before the Main Experiment started, each of the subjects who had so far volunteered to take part was asked to fill in a list of 50 objects which in his opinion could be drawn easily and unmistakably. It was stressed that these lists should be the individual work of each subject, and that as far as possible the lists should be compiled in the same room, and at the same time as the subject would normally be doing the experiment.

The lists were then collected and security measures taken to prevent any subject getting access to them. This material was then analysed by helpers, and the frequency that each object had been selected determined. The lists returned by the analysers originally listed 15,828 titles, 992 of which were "singles" --- that is, they were shown as occurring only once in all the suggestions analysed. It was originally intended to select only objects which had been suggested once to use as targets. For this purpose an attempt was made to compile a list of 100 titles as unlike one another as possible, so that no confusion of shape would result, and which would also not be subject to seasonal changes. Finding objects which satisfied these two criteria proved unexpectedly difficult, so that eventually some items were included in the list which were in the experimenter's view not completely ideal. From these 100 titles the 50 Experimental Target Originals, and the 50 Control Series Originals were selected by an essentially random method which will be described fully later. After the experiment had been completed, the experimenter checked through the Preliminary Frequency lists to find which subjects had suggested the titles actually forming part of the Control and Experimental Series. He found that seven of the 100 titles actually used occurred twice and not once in the original suggestions as the analysers had supposed. The correct total number of suggestions analysed is thus 15,835. Moreover, the inclusion of these seven "doubles"

may result in a slight initial inequality in difficulty between the Experimental and Control Series Originals. This initial inequality will have to be allowed for if necessary in the assessment of the results.

Thus the Preliminary Frequency Experiment was carried out for three reasons:

(a) To determine empirically how many hits could be expected on each original from chance alone.

(b) To reduce this chance factor to a minimum.

(c) To select originals that would be approximately equal in difficulty, so that later valid comparisons could be made between various parts of the experiment.

The writer has often been asked why he did not choose originals as targets which had not been suggested at all in the 15,835 suggestions returned by the subjects. If this had been done there would have been no way of knowing what the chance frequencies of thinking of these originals were, and it could not be assumed that they were equal in difficulty, but only that their chance frequencies were less than 1 in 15,835. Also infinitely small probabilities might be introduced into the calculations which would be difficult to deal with statistically.

Needless to say, the subjects were not told what use would be made of their Preliminary Frequency Lists, and they had no way of inferring from their own lists which of their suggestions might be used as targets in the Main Experiment, even if the possibility had occurred to them.

Careful precautions were also taken to prevent them from seeing any list but their own, or the frequency analysis carried out on the lists, or the list of 100 possible originals finally selected. It can safely be concluded that nothing in the Preliminary Frequency Experiment would be of any help to the subjects in enabling them to score hits in the Main Experiment.

A copy of the instructions issued to the subjects for this Preliminary Frequency Experiment is given overleaf:

SURNAME (in capitals):

INITIALS:

SERIAL No:

EXPERIMENTS IN PARANORMAL COGNITION: No 1.

PRELIMINARY FREQUENCY EXPERIMENT.

The aim of this preliminary experiment is to find out what kind of objects normally come into your mind under conditions similar to those of the main experiment. When you fill in this form, therefore, please do so:

- (a) In the room where you will probably be doing the main experiments.
(b) While you are alone, and without consulting anyone else as to what you should put down.
(c) At the time of the day when you will normally be doing the main experiments.

INSTRUCTIONS: Each day, for five days, jot down the names of ten objects or things in the spaces provided below. Use block letters.

The objects may be of any kind, provided that they can be drawn unmistakably and fairly easily. Think of simple objects and not complex scenes.

Allow an interval to elapse between the first word you write down, and the next. Fill this interval by doing something else, so that the words you write down will be unrelated, and not all part of one train of thought.

Table with 5 columns labeled 'Date:' and 10 rows numbered 1 to 10. Each cell contains a date from 11 to 50.

Fill in the date at the head of each column. After five days, when you have completed the list of fifty words, bring this sheet to the Psychology Department Library, and put it into the box labelled "Paranormal Cognition Data" which you will find on a table there.

In what room did you draw up this list ? (Give address and name or number of room, etc.)

.....

At what time of the day did you draw up this list ?.....
The main experiments will start at the beginning of Third Term.

(6) The selection of Control and Experimental Series Originals

The next step in the procedure was to divide these 100 possible originals into two series of 50 each, the one series to be the Experimental Target Originals, actually used in the experiment, the other to be the Control Originals. It was necessary that all subjective factors of selection be eliminated in allocating originals to one or other of these two series, and that the basis of selection be a completely chance and random one.

To accomplish this the titles of the 100 selected possibilities from the preliminary frequency lists were printed on squares of cardboard, about one inch square, and these 100 cards put into a wooden selector box, which was completely enclosed, and had a slit only large enough to allow one of the cards to be shaken out at a time.

This box was sent sealed to the agent. At the beginning of the experiment he broke the seal, and each day determined the two drawings he was to make by shaking two titles from the box, and then making two simple outline drawings of the objects so selected. He was warned to accept whatever two cards came first from the box, and not to exercise any powers of selection himself.

In all 50 such Experimental Series Originals were made by the agent during the course of the experiment, leaving 50 unused slips still in the box, which had not been seen by him, and which had taken no part in the experiment.

Some months after the experiment had been completed someone unconnected with the experiment was asked to make drawings of these titles similar to the drawings made by the agent for the Experimental Series Originals during the experiment. These 50 Control Series Originals were therefore in every way equivalent to the 50 Experimental Series Originals, except that they had not been used as targets in the experiment.

(7) Time sequence of Experiment:

The experiment was divided into five sections of five days each and took place as follows:

Section A: Monday to Friday, August 11th to 15th, 1947.

Section B: August 18th to 22nd, 1947.

Section C: August 25th to 29th, 1947.

Section D: Sept. 1st to 5th, 1947.

Section E: Sept. 15th to 19th, 1947.

This gives 25 experimental days in all.

(8) Procedure followed by the agent:

On each of these days the agent in Cape Town shook two cards from the selector box, made two separate drawings of these two objects, and locked these two targets in a light-tight box, where they remained until the following morning, when the next two targets had to be made. They were then removed from the exposure box, put into opaque envelopes, and locked away with the other material used by the agent. The agent kept all material relating to the experiment locked away. The subjects in Grahamstown were kept ignorant of the agent's name and address, so they had no means of knowing, in any normal way, what drawings he had made.

Since it is important to know exactly what procedure was adopted by the agent, and whether the precautions taken to rule out sensory leakage and rational inference were completely watertight, the instructions issued to the agent, and his own statement of the procedure adopted by him are given in full overleaf:

(9) Instructions issued to the agent:
and procedure adopted by him:

EXPERIMENTS IN PARANORMAL COGNITION.

Outline of general instructions and information
given to agent before the experiments started.

The agent was sent a summary of the aims, and an outline of the experimental procedure which would be adopted. He was also given a sample of the record booklets which the subjects would be using.

He was not told that he was the only agent making drawings, and during the experiment was under the impression that there was a second agent operating in Johannesburg.

For section D & E of the experiment he was given the results of his personality tests, and other information, and told to write a letter to his group of subjects presenting this material in his own style.

He was also instructed as follows on the precautions necessary:

"It is necessary in experiments of this kind to take elaborate precautions to rule out the possibility of subjects acquiring by any normal means knowledge of what the drawings are. Telepathy is so against established dogma that if positive results are obtained, critics will not hesitate to use even the most flimsy possibility of the subjects having obtained knowledge in normal ways to try and discredit the results. It is therefore necessary to take the most rigid precautions to prevent such criticism. Even though the subjects will be hundreds of miles away from Kenilworth, please observe these additional precautions:

(1) Keep the selector box, the diary, the drawings you make, and all other material used in the experiment locked up unless you are actually using the material.

(2) While the experiment is on, do not discuss it with anyone else, even your own family, and don't mention to anyone what you are doing. Nobody but you should see your drawings. As we are investigating the relationship between your mind and those of the subjects, if other people see the drawings uncontrolled factors are introduced.

(3) At the end of every section of the experiment seal up the drawings used for that section, your notes, and anything else in the envelopes provided. In all ways try to make each section of the experiment as distinct and cut off from the other sections as possible. The reason for this is that it seems as if E.S.P. doesn't take much account of time. Subjects are as likely to score hits on section "C" drawings before they are made as they are when the drawings are actually made. Perhaps we can prevent this by making the various sections of the experiment as distinct psychologically from each other as possible, and distinct also in space and time. If we can get the subjects scoring only on the current sections originals, we may be able to use the second type of validation outlined in the summary.

(4) At the end of the experiment please supply a signed statement of the precautions you have taken, and say whether you think it has been possible for anyone to obtain normal knowledge of the drawings you have made.

(5) Confine your attention as far as possible to the current day's drawings and other material. As far as possible do not look at the drawings or diary sheets which have already been completed.

(6) Note any relaxation of precautions or variation of procedure in the diary.

(7) Once you have shaken the title of an original from the selector box, draw that object. On no account return a slip to the box for any reason, or exercise any other process of selection. It is essential that the originals you are to draw should be determined purely by chance, and you should not reject any original as being too hard to draw etc.

(8) Do not worry if your drawings are not very good. In fact, as long as your drawing is recognisable it is O.K. Draw only the simple object indicated, and not the object in a setting.

(9) Amount of concentration required: The idea of going over the object in Indian ink is to get it firmly into your mind. What evidence there is seems to suggest that the agent doesn't have to be concentrating on the object for the subject to get the impression. The source of the subjects' impressions seems to be the agents sub-conscious mind. Hence I don't want you to concentrate especially throughout the day on the objects you have drawn. You will have the objects in the forefront of your thoughts:

- (a) When you are drawing them
- (b) When you are writing up the diary
- (c) When you remove them from the exposure box on the following morning.

There is the additional factor that some subjects may get their impressions direct from the drawings themselves, clairvoyantly, so that no special concentration is necessary.

(10) Detailed instructions will be sent to you later on the actual procedure to be followed in selecting and making your drawings."

EXPERIMENTS IN PARANORMAL COGNITION.

General instructions to the agent for
selecting originals and making drawings

On the day that the experiment starts (August 11th) between 7.00 and 7.30 a.m. break the seal over the slot of the selector box.

Shake out two squares of cardboard from the box. These squares will bear the titles of the two things to be drawn.

Lift one of the special drawing sheets provided away from the remainder, so that there is no chance of subsequent sheets bearing the marks of the drawing you are going to make.

Sketch the first object in pencil as well as you can on the sheet.

Ink over the outline of your sketch with a brush and indian ink.

Fill in the information required in the diary.

Repeat this procedure for the second drawing.

Lock the two drawings made in the special box provided.

Put the squares of card bearing the titles into the envelope marked "Section "A" Originals" and lock this envelope up.

Lock away the selector box.

Make sure that you have left no information around that would give a clue to the drawings in the box, or to the drawings that might still have to be made.

On the following day, between 7 and 7.30 a.m. remove the drawings from the box, slip them into the brown envelope for section "A" material, and lock this envelope away, after repeating the procedure outlined above to get the next two drawings.

As each section of the experiment is completed, seal all the material pertaining to it in the brown envelope provided.

N.B. Don't put anything on the drawing sheets which would give a clue to the order in which they were made, or when they were made, once the serial numbers have been covered up.

Use of coloured squares in sections B and E of experiment:
Follow the standard procedure outlined above in making your drawings. When you have finished inking them in, try to imagine what they would look like if drawn in the same colour as the square of card on the bottom of the sheet. Thus you will imagine them in a different colour each day.

EXPERIMENTS IN PARANORMAL COGNITION.

AGENT'S ORIGINAL
DRAWING NUMBER:-

AGENT'S ORIGINAL
DRAWING NUMBER:-

Nothing should appear on this sheet except the drawing whose serial number appears above, and its title below. All other details should go in the diary provided.

Print drawing's title here:

COUNTERFOIL:

(Do not write on
this counterfoil).



Serial No.:

STATEMENT BY THE AGENT ON THE
PROCEDURE ADOPTED BY HIM DURING THE EXPERIMENT.

TO ALL WHOM IT MAY CONCERN:

During the experiment in para-normal cognition, recently completed, in which I acted as agent, my general procedure was as follows:-

Each morning, between 7.10 and 7.15 a.m. (with one or two exceptions), I shook out two squares of cardboard from the selector box, each of which contained one of the titles of the drawings I was to make that day. I then got out of bed, took two clean sheets of paper from those supplied to me by the person carrying out the experiment, and made a rough pencil drawing of the objects (one on each sheet). When satisfied I went over the outline with Indian Ink.

I then removed the drawings made the previous day from the box, and in their place put the two new drawings; I placed the two drawings from the previous day in an envelope, and locked the envelope away in my cupboard. The two squares of card I put in another envelope, and also locked it away. I locked the box with the two drawings for the day and kept the key, together with the key of my cupboard, in my pocket throughout the day.

In the evening, just before retiring to bed, I completed the diary for the day, the first part of which I had filled in immediately after having made the drawings in the morning. This diary was kept locked away at all times when not being used.

During the period of the experiment, I took special care not to mention to anyone that I was even doing the experiment, and certainly never mentioned what the particular drawings for the day were. I did not even tell the members of the household that I was doing the experiment, and even if any of them realised that I was doing it, they could certainly not have found out what the drawings on any particular day were.

There was thus no possibility of any person having prior knowledge of what the drawings were at all, and certainly no attempts were made to transmit any information to anybody anywhere.

In addition to all these precautions, I purposely refrained from writing to the experimenter while the experiment was actually being carried out.

Signed by the agent.

The original of this document is filed with the material of the experiment, and can be seen upon application.

(10) Procedure followed by the experimenter and

Subjects in Grahamstown:

The procedure followed in dealing with the subjects was as follows. On the Sunday just prior to the beginning of each section, each subject was issued with a book in which the ten drawings for the next week were to be made, together with the general and specific instructions they were to follow, and any linkage material. At the end of each section, on Saturday, the subjects' books were collected from their rooms. Care was taken to ensure that all the books issued for one section had been collected, returned or destroyed before the books for any subsequent section were issued. Officials were appointed in each residence to collect and distribute the test books personally. This was done in an attempt to discourage subjects from inspecting the drawings made by other people taking part in the experiment. It was also thought that less subjects would drop out if they were not required to collect and return their books themselves, but instead had the books delivered to their rooms. It was also found that some subjects had to be constantly reminded and stimulated to produce their drawings, and the frequent personal contact of the distributing officials and the experimenter with such delinquents tended to keep them up to the mark.

As a matter of policy the experimenter also tried to see as many of the subjects personally each week as possible, to keep up their interest in the experiment, and prevent any subjects dropping out who could possibly be persuaded to continue. Since they were never told what the experimental originals were, nor whether they were succeeding or failing, it was easy for them to lose interest. Steps were taken to supply any mildly ill subjects in the sanatorium with the necessary books, but no attempt was made to force gravely ill subjects to complete the experiment. These measures resulted in a total of 17,440 drawings being returned by the

subjects out of a possible 18,550, which is a 94% return.

In their books the subjects were asked to try and reproduce the two drawings made by the agent each day. They were asked as far as possible to make their drawings in the same room, and at the same time each day.

(11) General instructions issued to the Subjects:

A copy of the general instructions issued to the subjects, and a sample of the blank sheets on which they made their drawings are given overleaf. The general instructions were issued unchanged for each section of the experiment. In addition, the subjects each received a sheet of specific instructions which related only to the section of the experiment which was in progress at the time.

The subjects' sheets on which they made their drawings were slightly larger than the sample shown. They occupied a whole page of foolscap. The linkage squares of colour were normally placed in the same position as the square on the sample sheet.

EXPERIMENTS IN PARANORMAL COGNITION.
EXPERIMENT II.

Subject's name:		Initials:
Address:.....		
Serial Nos:	to	Group:

GENERAL INSTRUCTIONS:

Please read these instructions carefully before starting the experiment.

EXPERIMENTAL SCHEDULE:

The experiment has been divided into five sections, and will take place as follows:-

Section "A":	Monday August 11th.	to Friday 15th.	inclusive.
"B"	18th.		22nd.
"C"	25th.		29th.
"D"	Sept. 1st		5th.
"E"	15th.		19th.

There is thus a break of two days between each of sections "A" "B" "C" and "D", and a break of nine days between sections "D" and "E".

HOW THE TARGET DRAWINGS WILL BE MADE:

Each day, while the experiment lasts, your agent will make two different drawings. These drawings will NOT be of complex scenes, but of simple objects that can be drawn unmistakably and fairly easily.

Each pair of drawings will be made at 8 a.m., and left in position in the town where the agent lives for 24 hours until 8 a.m. the following day.

WHAT YOU HAVE TO DO:

You are asked to try and reproduce, in this book, the two drawings made by your agent each day. Since these drawings will be made in a different town, you will have no normal way of knowing what they are, and will have to rely on getting an impression of them by mental telepathy.

The results of previous investigations suggest that it is better NOT to concentrate intensely when trying to get such impressions, but rather to adopt a completely passive attitude; withdraw your thoughts from your immediate surroundings, and wait for suitable objects to come to mind.

Naturally you will not know whether the drawings you make in this way resemble those of your agent or not. You may

feel confident that they do, or you may feel no confidence in them whatsoever.

KEEP THE CONDITIONS STANDARD:

You may, if you wish, adopt a different approach from that outlined above, but whatever method you decide on should be used unaltered throughout all sections of the experiment.

Similarly, although you can attempt to reproduce your agent's drawings at any time while they are in position, keep as far as possible to the same time each day, and work in the same room.

LEAVE NO BLANKS:

In order to be of use for the purpose of the experiment it is essential that two drawings be made by you on each day of the experiment. Please make an attempt of some kind at each drawing, whether or not you feel that you are succeeding.

PIN UP THE REMINDER SLIP:

As a reminder a printed slip reading "Have you remembered your TELEPATHY DRAWINGS today?" has been provided. Tear this slip off, and pin it up in your room so that it will catch your eye each day.

HOW TO DEAL WITH OMISSIONS:

If despite this precaution you find that you have omitted to do some drawings on the appropriate day, do NOT leave the spaces blank, but fill them in as soon as you discover your error. Indicate on the counterfoil when the drawings were finally made.

WHERE TO MAKE THE DRAWINGS:

Please make all drawings in this book, NOT on rough paper first. Drawings should not be altered or "finished", and it doesn't matter if you cannot draw well. Print the name or description of the thing you have drawn at the bottom of each sheet. Do not write anything else in the spaces reserved for the drawings, nor make any note on them that would give a clue to the date or order in which they were made once they have been detached from the counterfoils.

If you make more than one attempt at a drawing, write "Best" next to the one you think most likely to be successful.

WHERE TO MAKE ADDITIONAL COMMENTS:

If you want to make any comments on your drawings, do so in the space labelled "Notes and Impressions" on the counterfoil. This space is not intended for elaborate introspections, but for verbal recording of any "ideas" that are too difficult to draw, and to give you an opportunity to make any comments you feel might be relevant. There is no need to make any entry here unless you wish to do so.

FILL IN THE COUNTERFOIL:

Please, however, fill in the remaining information asked for on the counterfoil. Under "Physical and Mental Condition" note briefly whether you were feeling normal, ill, tired, excited, worried, etc., when you made the drawings. It has been found that telepathic ability is also affected by various drugs, so please give your approximate consumption each day up to the time of making the drawings.

Write "+" after "Confidence" if you felt you were getting a genuine impression of the target drawing made by the agent, and "o" if you felt that you were not.

Fill in the date and time of making each drawing on the counterfoil, and indicate under "Clear Image?" whether you had a clear picture in your mind's eye while drawing.

USE THE SHEETS IN THEIR CORRECT ORDER:

Be careful to use the sheets in the order which they occur in the book, and on NO account detach your drawings from their counterfoils.

INSERT CARD:

Please insert the card provided between the sheet you are using and the following ones, so as to prevent the possibility of marks made on one sheet showing on the others.

WORK ALONE:

It is desirable that reproductions by the various participants should be quite independent of each other, so please do the experiment while alone, and do not discuss the drawings you make with other subjects until after the whole experiment has been completed, that is, not until after September the 19th.

CHECK NAME AND INITIALS:

Make sure that your correct name and initials appear on the cover of this book.

DISPOSAL OF THIS BOOK:

Subjects not in residence at R.U.C. are asked to bring or post their books, as each section is completed, to: Mr. M.C. Marsh, c/o Psychology Department, Rhodes University College, Grahamstown.

If you are in residence at Rhodes, this book will be collected from your room when completed, and the book for the next section left in its place. If you are due to start a section, and have not received the necessary book, please contact Mr. Marsh immediately.

This sheet has been slightly reduced in size, to allow a binding margin here.

Normally it occupies the whole sheet.

No: No:
Time..... Time.....
Confidence..... Confidence.....
Clear Image ? ... Clear Image ?.....

Only a drawing and its name

are to go on this sheet

Only a drawing and its name

are to go on this sheet.

Physical and mental condition:

Consumption of drugs since getting up till time of making drawings:

Cups of tea..... Of coffee.....

Cigarettes.....

Alcohol...(Specify).....

Other Drugs: (i.e. Aspirin etc.)
Specify:

Notes and Impressions:
(Optional.)



COUNTERFOIL.

Print here the name of the thing drawn above.

Print here the name of the thing drawn above.

(12) The use of Control and Experimental Groups of Subjects:

In order to investigate the effects of the linkage material supplied to them, the subjects were divided into two groups, approximately equal in age, sex, etc. The principle was adopted of making one complete residence either all Control, or all Experimental subjects. It was thought that this would lessen the temptation for any subject to try and find out what linkage material was being supplied to the group to which he did not belong. These two groups were told that they would not have the same agent, but would each have a different agent living in Cape Town. They have never been told that in reality there was only one agent, and that one group was being used merely as a control.

Throughout the experiment both groups were given precisely similar instructions, but whereas the Experimental Group Subjects were given correct linkage material, the linkage material given to the Control Group Subjects was incorrect, and had no reference to the agent at all. Care was taken to make the two sets of linkage material as comparable as possible both in quantity and quality, the only difference being that the linkage material supplied to the Experimental Group Subjects was genuine, whereas that supplied to the Control Group Subjects was false.

No linkage material was provided for section A of the experiment, in order that this section might provide an indication of the normal scoring level of the two groups of subjects when no linkage was operating. This would show whether the selection method had succeeded in equating two groups in terms of basic E.S.P. ability. Once it had been demonstrated that the two groups were in fact equal in E.S.P. ability at the beginning of the experiment, any subsequent differences in scoring rate which they showed in subsequent sections could then be attributed to the new factors of linkage which had been introduced.

Thus, if the results of the Experimental Group of subjects showed an improvement when linkage was introduced, whereas the results of the Control Group Subjects did not change significantly from the level they had established in section A, this would suggest that the correct linkage material which the Experimental Group Subjects had been given was the factor responsible for their improved scores.

If, however, both Control and Experimental Group Subjects improved significantly on their section A scoring level, but showed no significant difference between the two groups, this would suggest that the virtue of the linkage material lay in its suggestion value only, as both correct and incorrect linkage produced the same effect.

If both groups improved significantly on their section A performance, and in addition the Experimental Group Subjects did significantly better than their Control Group counterparts, this would suggest that the virtue of the linkage material lay both in its suggestion value, and the factor of its correctness.

Finally, if neither group showed any significant improvement on their section A scores, this would indicate that the linkage introduced in this experiment had not been effective in increasing the scoring ability of the subjects, and that in this special instance at least, the validity of Carington's theory, and of psychometry had not been upheld.

It might possibly be argued that any improvement shown by the two groups might be due merely to the passage of time: that it merely meant that both groups of subjects were warming up. However, all the weight of the experimental evidence is against such an effect. Usually investigators report a decline in scoring level as an experiment progresses. Thus, if an improvement does occur, it would be legitimate to relate it to the additional experimental factors that had been introduced, and not to a time factor.

After these general remarks, we can now turn to an examination of the specific procedure adopted for each of the five experimental sections.

(13) The linkage factors introduced in the various sections of the experiment:

(a) Section A of the Experiment:

For this section of the experiment both Control and Experimental Group Subjects were merely told that their agent was carrying out the experiment in Cape Town, and that this first section was intended merely to establish what scores could be expected from the groups when very few linkage factors were operating.

(b) Section B of the Experiment:

In this section an attempt was made to provide a specific K idea for each pair of drawings made by the agent, since if Carington's theory is correct, one would expect this would increase the subjects' scoring capacity by making the essential bonds more distinct and numerous.

On each drawing sheet used by the agent was pasted a distinctively coloured square of cardboard. After he had made his drawing the agent was asked to associate it with the square of colour by trying to imagine what the drawing would look like if it had been made in the same coloured ink as the square.

The Experimental Group of Subjects were given on each page of their books similar squares of colour which corresponded exactly to those being used by their agent that day. Before making their drawings they were asked to sit down and fix their attention on the square of colour. This procedure made sure that the same K idea was present in the minds of both the subjects and the agent, and that the agent had definitely consciously associated the K idea with the drawing he had made.

A great deal of thought was given to the choice of suitable K ideas. Thirty of these were needed, and it was necessary for them to be sufficiently unlike one another not to lead to any confusion. At the same time they had to be non-specific in character: that is, they should not suggest any particular objects to the subjects. The use of inkblots was considered and rejected because it was felt that they would be too structured, and would allow too much opportunity for projecting unconscious material, as well as being difficult and expensive to reproduce.

Although colours were finally chosen, their use was far from ideal, as it was found that they tended to produce "tied" associations in the subjects' minds. Thus a red square tended to call out objects such as blood, post box, etc, constraining the subjects to these objects when they came to make their drawings.

Many subjects believed that their agent was making a drawing of an object whose normal colour was that of the square of colour on their sheets, and as the instructions did not explicitly state that the colour had no direct connection with the type of object drawn, they tended to avoid drawing objects that were not of that colour.

The Control Group Subjects also received squares of colour, but these were of course incorrect, and did not resemble those being used by the agent except in actual size.

The following colours were used: .

THE LINKAGE COLOURS USED IN SECTION B
OF THE EXPERIMENT:

Colours appearing
on the Agent's
Original Sheets,
and on the sheets
of the Experimental
Group Subjects:

Colours appearing
only on the sheets
of the Control
Group Subjects:

Monday,
August 18th:



Tuesday,
August 19th:



Wednesday,
August 20th:



Thursday,
August 21st:



Friday,
August 22nd:



(c) Section C of the Experiment:

This section was designed to test the effect of giving the subjects an object that had once belonged to, and been handled by the agent, to hold while ~~making~~ their drawings. It set out to test the value of psychometry, in comparison with other techniques.

Prior to the experiment a number of handkerchiefs were obtained from the agent which he had been using for a number of months. These were cut up into small squares. Each member of the Experimental Group of subjects was given one of these squares of material with the specific instructions issued for Section C. They were told to hold this material in their hands while trying to get their impressions. These squares of handkerchief were handled only by the experimenter after they had left the possession of the agent, and care was taken to handle them as little as possible. An example of these squares of material is given below:



The Control Group Subjects were given similar squares of handkerchief which had been the property of someone wholly unconnected with the experiment, and which had been handled only by the experimenter since leaving that person's possession.

(d) Section D of the Experiment:

The aim of this section was to try and find out whether emotional closeness had any effect on the E.S.P. scoring level. A common belief is that telepathy is more likely to occur between people who know each other very well, and who are linked by emotional bonds, than between total strangers. There is some experimental evidence to suggest that such relationships do have an effect on the telepathic process. (49)

Consequently the linkage material supplied to the subjects in this section of the experiment was designed to break down the ignorance of the subjects about their agent, and if possible make them feel that they knew him quite well. The Experimental Group Subjects were thus given an actual photograph of their agent, taken in his room at Cape Town, and showing him sitting next to the exposure box in which he put the target drawings each day.

In addition, they were told some of his personality test results (prior to the experiment he had completed the same personality tests as the subjects), and they were also supplied with a friendly letter written by the agent, telling them some of his intimate personal history.

The Control Group Subjects were given material quantitatively and qualitatively equal to that received by the Experimental Group Subjects, but in their case the photograph was of a man wholly unconnected with the experiment, who lived in Johannesburg, the personality test results were fictitious and referred to no living person, and the personal history was that of a man wholly unconnected with the experiment, living in Australia.

Before making their drawings each day, both groups were asked to read through the information booklets issued to them, and to have the photograph of the agent before them when they began to make their drawings on any particular day.

Copies of these information booklets are given overleaf.

EXPERIMENTS IN PARANORMAL COGNITION: EXPERIMENT II.DESCRIPTION OF AGENT FOR SECTIONS "D" AND "E".

Subject's Name:	Initials:	Subject's Group:
--------------------	-----------	---------------------

Address:

Booklet's Serial No:

<p style="text-align: center;"><u>WARNING</u></p> <p>This booklet should only be seen by the subject whose name appears above.</p> <p>The information contained herein is confidential and must on no account be made public.</p> <p>It comprises a photograph and a personality sketch of the agent serving the group indicated above.</p>

N.B. Please return this booklet with your drawing book at the end of the week.

CONFIDENTIAL INFORMATION ON THE AGENT
SERVING GROUP "E" SUBJECTS.

THIS IS A PHOTOGRAPH
OF YOUR AGENT, AND
OF THE BOX HE LOCKS
HIS DRAWINGS IN EACH
DAY AFTER HE HAS
MADE THEM.

NOTICE PARTICULARLY
THAT THE BOX IS
MARKED WITH A CHECK-
ERED PATTERN.



HOW HE SELECTS AND MAKES HIS DRAWINGS:

He has been provided with the names of a large number of objects, real and imaginary, mythical and commonplace, which can be drawn fairly easily. Each morning he selects two objects from these by chance, so that his choice is a random one, and he has no hand in deciding what he is going to draw. Having obtained the names of the two things he is going to draw, he first sketches them roughly in pencil on the special sheets provided, then inks them in using black Indian ink. He puts them in the checkered box shown above in the position occupied by the blank paper, shuts and locks the box, and leaves the drawings in position until the following morning.

ABOUT YOUR AGENT:

He has been persuaded to write the following personality sketch of himself for you:

"..... He says that he is very pleased to make this introduction to his group of subjects, mainly because he finds it always very interesting to talk about himself, (although he fears that others might not think the same,) but also because a little bit of his heart is at Rhodes.

He was born in Pretoria in 1926, and so recently celebrated his twenty-first birthday with a large and successful party. He started off by going to Pretoria Boys' High, (which he still considers a very good school,) but he feels that his real Alma Mater is St. Andrew's School, Bloemfontein. In this he is somewhat moved by the thought that that school has produced Lindsay Tuckett, who recently was rather a success as a batsman in the Fifth Test.

From this it may be deduced that your agent is very interested in Cricket: he plays it enthusiastically but, he fears, not always successfully. In addition, he is a fanatical Rugby-fan, but he prefers sitting on the touchline to playing it. He is apt to get wildly excited at matches where the University of Cape Town is playing. In addition, (if it can be counted for anything,) he wins a great deal of money at Bridge, but he is

not sure whether this is because he is a good Bridge player, or merely because he is unlucky in love.

He is considered by some to have crammed a great deal of experience into his short years: he has worked in no fewer than two different Government Departments, with a total of five different offices in various parts of the country. He also served in the S.A.A.F. as an air-gunner, spending too much time in Egypt and too little in Italy. He did some of his training at Port Alfred, and twice visited Grahamstown, (and a very excellent pub. there,) but what he saw of it he liked. He visited the Cathedral there and was duly impressed.

He is somewhat dubious about letting you know the results of his personality tests, and sympathizes with all you who write them also, but he is going to throw himself on your mercy and tell you what he is like, more or less.

His three main wishes at present are, firstly for a large, well-bound library, with plenty of classics and Latin Texts in it. Secondly, for a comfortable house with good service, and thirdly for a large wardrobe.

Other wishes are to go to Oxford; to have plenty of money so as to be able to travel all over Europe; plenty of leisure time to spend in libraries and museums, and to be able to write detective stories in the style of Michael Innes and Edmund Crispin, (whom he considers to be one and the same person.) He would like to belong to the best Conservative clubs in London, and visit them at various times so as to be able to meet the best people, and from time to time, do the 'Times' crossword puzzle.

As far as his suppressed desires go, he says that if you are really curious about them you will have to find them out telepathically, which you will probably be able to do if you try very hard.

His main interest in life is his Aesthetic one, which means that he is interested in artistic things, especially music, (he likes Beethoven most of all,) and that he values each bit of experience in life for its own sake. In a group of 185 Rhodians only one man and five women have a higher artistic value.

Besides this artistic interest, one which is almost as important, (according to the tests,) is his desire to control and organise people, only two men and one woman having a stronger political value than his in the group. He is not interested in being with people for their own sake ---- no Rhodian in the group has a lower social value.

He is also, according to the tests, not interested in formal religion; only 14 out of 185 Rhodians have a lower religious value.

As far as the scientific and economic things of life go, his interest is about average.

In answering his personality questionnaires he was more truthful than 60 % of the Rhodians, and from the results appears about as happy as most people at R.U.C., feeling himself at one with his group's ideals and attitudes to about the same extent.

At the moment, however, he is not prepared to be very sympathetic towards other peoples' troubles, only ten Rhodians, (all men) being more unsympathetic.

On the other hand, he is more conscious of a main purpose in his life than most Rhodians, and has organised his life better than they have to achieve his aims.

Although he is inclined to be impulsive in his personal relationships at times, in non-personal matters he relies more on his judgement. He is also slightly more controlled when acting under emotional stress than the average Rhodian, and more stable as well.

Although he is neither particularly introverted nor extraverted, and only as confident as the average Rhodian, only 21 of the 185 Rhodians would be able to dominate him in face-to-face situations.

He thinks by now that you know him only too well. At any rate he hopes that the information he has given you will be of some assistance to you in the experiment....."

TO BE FILLED IN BY THE SUBJECT:

If you recognise your agent from the above description, write his name here:

How well do you know him ?

How well do you know Kenilworth, Cape Town, where he lives ?

N.B. The personality test results of the subjects who have completed their personality tests will be sent out with the drawing books for section "E" of the experiment, on Sunday September 14th.

CONFIDENTIAL INFORMATION ON THE AGENT
SERVING GROUP "C" SUBJECTS.



C

THIS IS A PHOTOGRAPH OF YOUR AGENT AND OF THE BOX HE EXPOSES HIS DRAWINGS IN EACH DAY AFTER HE HAS MADE THEM.

NOTICE PARTICULARLY THAT THE BOX IS PERFECTLY PLAIN, AND BEARS NO SPECIAL MARKINGS OF ANY SORT.

HOW HE SELECTS AND MAKES HIS DRAWINGS:

He has been provided with the names of a large number of objects, real and imaginary, mythical and commonplace, which can be drawn fairly easily. Each morning he selects two objects from these by chance, so that his choice is a completely random one, and he has no hand in deciding what he is to draw. Having obtained the names of the two things he is going to draw, he sketches them in white chalk on black cardboard sheets. He puts them in the plain box shown above in the position occupied by the blank paper in the photograph above, and leaves them exposed until the following morning.

ABOUT YOUR AGENT:

He has been persuaded to write the following personality sketch of himself for you :

"..... To present this "personality picture" of myself to you I must view myself in retrospect over a period of four decades. I was the youngest of four children, my father being a successful businessman in a small country town. Both my parents were intensely religious and strict, so that during the pre-adolescent period of life, I, along with my brothers and sisters, was really a "model" child in my acceptance of my life of religious experience.

I was an extremely healthy young fellow in my boyhood, and very active physically. I attended school in Port Elizabeth. In the course of my study to matric., I topped my class in English, History and Maths every year, gaining a University Scholarship for an Arts course.

My chief intellectual interests at school had been literary, and quite creative. I had written about 40 sonnets of rather uneven quality, several ballads, and a great quantity of verse, most of it in romantic strain; some one act plays, and a few hundred lines of blank verse ---- "sonorous and majestic" as a teacher ironically described it.

I excelled in comparison with my schoolfellows in such things as putting the shot and throwing the hammer, and swam a very rhythmic and successful breast stroke in representative teams.

Good music has always fascinated me, and by 18 I had reached a pretty high technical standard. Musical forms I found most interesting, (possibly because of my mathematical bent.) Later I experimented in musical composition, but with only fair success. My products were of effort rather than inspiration. To those things I have mentioned add a passionate fondness for nature, and an increasing appetite for bush lore and the "open spaces".

You will note that as a young adolescent I would be considered a person of high promise.

At the end of my school days came the question of a career, and whether by chance or by as yet unfathomed design I chose teaching, and spent four years at an University Teachers' Training College, passing with high distinction. The College held me on its staff for two years rather than let me go to teach in schools, and my Professor of English continually stirred me with the prospects of an academic career.

I had very little money to spend at any time. This, and an apparant lack of interest in the fair sex, had deterred me from forming any emotional attachment to any particular female, though I was friendly with many as students, especially if their interests were literary, musical or dramatic.

As I look back I see two events which have been of major importance for my later life ---- my marriage, and the depression of 1930-1933.

When I met "Her" I was caught on the sky-high lure of romance, and after a whirlwind courtship we were married about three months after our first meeting. Our chief common interest was our fondness for out-door life. In the exhilaration of our excursions we felt that this was the right basis for marriage, but as it proved later to be, we were mistaken, though we lived our first six years in perfect happiness, and had two children, a girl, now sixteen, and a boy, now twelve.

Circumstances had necessitated me transferring from my College job to the School Teaching Service. I only later realised how great a blow to my career this transfer had been. This realisation came when my family responsibilities began accumulating though my salary cheques were not increasing as fast. I began to realise two things:-

(1) That my ambition to be a Literary Scholar and University Lecturer was likely to be permanently frustrated, and

(2) That at a time when I should have seen myself in true perspective and sifted out my plans, I had been blinded, if not temporarily paralysed by romance.

From this point in my life I seem to be able to trace a growing antipathy towards my wife and children. My music too had gone by the board, likewise my writing, and I had begun to seek, more and more, the company of other men. I became what is commonly known as "A Jolly Good Fellow" revelling in yarns over a stream of raucous laughter, enjoying risqué stories, and even baldly lewd jokes ---- but at heart really sad, disconsolate and perplexed. I saw no good in anything. Our marriage was rapidly leading to disaster, when suddenly, it seemed almost by an act of Providence, I was offered a particularly good job at a private school in Cape Town. I accepted, but the irony of it ---- three days after I had accepted I was offered a senior post at my old University, with all the things I had hoped for included in it, but the Headmaster of the Cape Town school held me to my contract, and I had to refuse it.

The next eight years saw me deputy headmaster in this school ---- but what of my marriage, you ask? My wife left me four years ago, and I carried the family as well as the school, and suffered the censure of my children.

In my desolation I remarried two years ago --- a widow with one daughter aged ten, and at the beginning of this year succeeded to the headmastership of the school. To all outward appearances I am a successful man. I leave you to read between the lines.

I had considered adding something about my personality test results. My main wishes and suppressed desires you will have gathered from what I have written above. I find, however, that my test results mainly reflect what I have written. I am told that I am more stable than 95 % of the Rhodes students, and more truthful and accurate in answering my questionnaires than 80 % of you. You may find these two facts of interest.

P.S. And now after collecting these papers, I ask myself (note how highly interested I am) "Why the devil should I have any soul in a psychology experiment ---- telepathy of all things?" ---- I wonder

TO BE FILLED IN BY THE SUBJECT:

If you recognise your agent from the above description, write his name here:

How well do you know him ?

How well do you know Kenilworth, Cape Town, where he lives ?

N.B. The personality test results of the subjects who have completed their personality tests will be sent out with the drawing books for section "E" of the experiment, on Sunday September 14th.

(e) Section E of the Experiment:

For this section of the experiment it was originally intended to combine all the linkage material previously issued for sections B, C and D. After section B, however, many subjects complained that the squares of colour supplied as linkage material had hindered their getting impressions. Consequently this original plan was abandoned, and it was decided not to issue them with colours or handkerchiefs again for section E. The agent did, however, repeat the procedure he had used in section B, and associated each of his drawings with squares of colour previously supplied to him. These squares of colour are reproduced on page 68.

The information booklets supplied in section D were returned to the subjects for section E, and they were asked to use them in the same way as they had done previously. In order to add fresh interest to this material, the subjects were also given their own personality test results, and it was suggested to them that they should compare their scores with those of their agent.

After section E had been completed, the subjects were asked to fill in and return the "Psychic Questionnaire". The administration of this questionnaire constituted the last stage of the experimental procedure, and after it no further data was obtained.

SUBJECTS' PERSONALITY TEST RESULTS
SUPPLIED TO THEM WITH THE LINKAGE MATERIAL FOR
SECTION "E".

To lend fresh interest to the personality material on the agent, which had been issued for section "D" of the experiment, and re-issued for section "E", the experimenter supplied each subject with the results of his or her personality tests. This was accompanied by a letter thanking the subject for taking part in the experiment.

Both letter and personality test material are given below.

Letter of Thanks:

Dear Subject,

As this is the last section of the experiment I want to thank you very much for taking part. I appreciate your kindness all the more because I know how difficult and boring it becomes to carry out such a routine task day after day. You may also have had to put up with a good deal of ridicule merely for your open-mindedness in taking part. If so, I'm sorry. Telepathy is a topic which arouses strong emotional attitudes in people.

Perhaps as some slight return for your participation you might care to have the results of your personality tests. They are given on the other sheet in this envelope, and I have tried to make them self-explanatory. One side gives your test scores, and the other side shows what these scores mean. If you would like to discuss any points in connection with these results with me, you can usually find me in the M.A. lab. of the Psychology Department (next to the typing room) in the mornings, and I'd be very pleased to help in any way I can.

Assessing the drawings you have made is going to take a long time ---- there are over 16,000 of them! I will, however, try to let you know the results of the experiment as soon as I can, probably through the medium of a lecture.

I would have liked to convey my thanks to you personally, but as that isn't possible I hope you won't mind them expressed in this form instead.

Sincerely,

M.C. Marsh.

Personality Test Results:

Subject's
name:

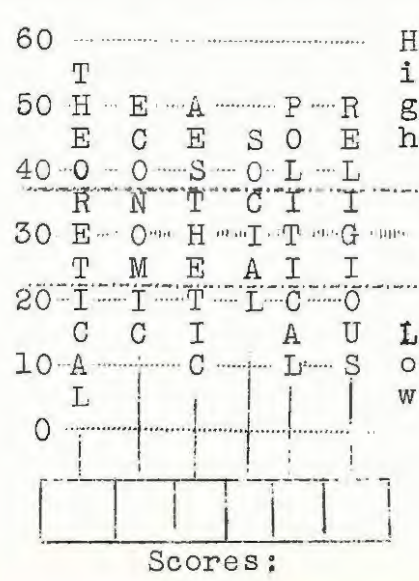
In these tables your score for each trait has been ringed. Suppose on looking down the "Accuracy" column you found "20-23" ringed. This would mean that your test score for this trait was 20-23. The figure "15" in the "No." column indicates that 15 out of the 340 subjects got the same accuracy score as you did. If you wanted to find out how many subjects were more accurate than you, you would add the figures appearing above the "15": (27+36+58+99+92 = 312) To find the number more inaccurate than you, you would add the figures appearing below the 15: (6+3+1+2+1 = 13)

Caution: Although the results of personality tests are accurate enough for comparing groups of people, taken individually the results may not give a completely true picture of the subject's personality. The test results should be taken as indications only, and not regarded as completely accurate.

PROVISIONAL DISTRIBUTIONS SHOWING THE NUMBER OF SUBJECTS REACHING EACH TEST SCORE ON THE PERSONALITY TRAITS LISTED BELOW.
(GROUP COMPRISING 340 SUBJECTS.)

ACCURATE		HAPPY		IDENTIFIED WITH GROUP		SYM-PATHETIC		PURPOSE-FUL		NOT IMPULSIVE	
Score	No	Score	No	Score	No	Score	No	Score	No	Score	No
0-3	92	0-1	146	0-5	52	0-3	39	18-23	2	0-1	114
4-7	99	2-3	23	6-11	51	4-7	40	24-29	19	2-3	37
8-11	58	4-5	43	12-17	66	8-11	50	30-35	42	4-5	82
12-15	36	6-7	31	18-23	64	12-15	55	36-41	62	6-7	33
16-19	27	8-9	17	24-29	40	16-19	48	42-47	59	8-9	30
20-23	15	10-11	31	30-35	19	20-23	41	48-53	64	10-11	20
24-27	6	12-13	12	36-41	24	24-27	30	54-59	34	12-13	8
28-31	3	14-15	8	42-47	7	28-31	18	60-65	31	14-15	10
32-35	1	16-17	10	48-53	11	32-35	8	66-71	11	16-17	5
36-39	2	18-19	9	54-59	2	36-39	8	72-77	10	18-19	0
Over	1	20-21	2	60-65	1	40-43	2	78-83	5	19-23	0
		22-23	2	66-71	1	44-47	0	84-89	0	24-25	1
		24-25	3	Over	2	48-51	1	90-95	1		
		Over	3								
IN-ACCURATE		UNHAPPY		ALIENATED		UNSYMPATHETIC		DRIFTING		IMPULSIVE	

CONTROLS EMOTION:		ADJUSTED		SUBJECTS SCORE:		No: J		No: S		No: I		No: D		No: S		No: N	
Score	No	Score	No	Plus:		U	M	P	Y	E	L	F	S	U	F	R	O
0-3	43	30-44	1	220 239													
4-7	55	45-49	2	200 219													
8-11	51	60-74	31	180 199													
12-15	64	75-89	55	160 179	2												
16-19	38	90-104	56	140 159	8												
20-23	28	105-119	61	120 139	10												
24-27	23	120-134	60	100 119	10												
28-31	22	135-149	23	80 99	17												
32-35	10	150-164	16	60 79	13												
36-39	3	165-179	20	40 59	21												
40-43	3	180-194	7	20 39	28												
		195-209	4	+0 +19	28												
		210-224	1														
UN CONTROLLED.		Over	3	-1 -20	25												
		MAL ADJUSTED		21 40	31												
				41 60	31												
				61 80	23												
				81 100	23												
60				101 120	28	S											
T				121 140	16	T											
50	H	E	A	141 160	15	A											
	E	C	E	161 180	5	B											
	O	O	S	181 200	4	L											
	R	N	T	201 220	2	E											
	E	O	H	Minus													
	T	M	E														
	I	I	T														
	C	C	I														
	A	C	L														
	L																
0																	



The "Study of Values" test opposite measures the relative importance in your life of six main values. It does not tell you how the strengths of your values compare with those of other people but merely what their relative strengths are in you. Scores of from 28 to 31 are average.

Interpretation of the Personality
Test Scores given above:

Note first that all these scores only tell you what you are like compared with other Rhodians. Even the person falling at the extreme of the Rhodian group might still be average in a non-Rhodian sample of the population.

"Accuracy": The first score given overleaf deals with your accuracy in answering the test questions. A high "accuracy" score, falling towards the "inaccurate" end of the scale indicates either that you did not pay sufficient attention to the wording of the questions, or else that you are inclined to bluff yourself or others as to what you are really like.

"Happiness": gives a measure of your happiness in your group here at Rhodes.

"Alienation": indicates whether you feel yourself at one with your group's aims and ideals, or whether you feel very different from the people round you.

"Sympathy": your "sympathy" score tells you whether you are sympathetic towards other people, or whether you are not very concerned with their troubles.

"Purpose": indicates whether you are conscious of any main goals in your life, and whether you have organised your life to achieve them, or are merely drifting.

"Impulsive": if your score in the next column falls near the "impulsive" end of the scale, this means that you live for the moment, and are not inclined to defer present benefits even though by so doing you would benefit more in the future.

"Control": refers to the way you express your emotions. If you allow them to influence your behaviour your score will be towards the "uncontrolled" end of the scale.

"Adjustment": the last column gives your general adjustment on all these traits combined, compared with the Rhodes group.

"Jumpy-stable": This first scale in the second set of scores is a measure of emotional stability. Persons scoring high on this scale tend to be emotionally unstable. Those scoring low tend to be very well balanced emotionally.

"Self-sufficient": persons scoring high on the "self-sufficiency" scale prefer to be alone, rarely ask for advice or encouragement, and tend to ignore the sympathy of others. Those scoring low dislike solitude and often seek advice and encouragement.

"Introversion": persons scoring high on the "introversion-extraversion" scale tend to be introverted; that is, they are imaginative and tend to live within themselves. Those scoring low are extraverted; that is, they seldom worry, rarely suffer emotional upsets, and seldom substitute day-dreaming for action.

"Dominance-submission": persons scoring high on the "dominance-submission" scale tend to dominate others in face-to-face situations. Those scoring low tend to be submissive.

"Self-conscious-Confident": persons scoring high on this scale tend to be hamperingly self-conscious and to have feelings of inferiority. Those scoring low tend to be wholesomely self-confident, and to be very well adjusted to their environment.

"Non-social-Gregarious": The final measure in this series. Persons scoring high on this scale tend to be non-social, solitary and independent. Those scoring low tend to be sociable and gregarious.

The third profile shown at the bottom left-hand side of the page deals with your main values in life.

Theoretical: The main interest of the person whose theoretical value is high is the discovery of facts, and the relating of these facts to the established body of knowledge. Usually scientists and research workers show a high score here.

Economic: A high economic value indicates preoccupation with practical considerations in life. Things are important to this type of person for their utility and money value. Businessmen and technicians fall into this group.

Aesthetic: A high aesthetic value indicates a person who appreciates each bit of experience in life for its own sake. This is a trait of imaginative workers like composers, artists and poets.

Social: Those with a high social value like to be with people for their own sake, and are unselfish and gregarious. Social workers generally fall into this group. The social type is not concerned with what he can get out of people, but what he can give. He or she enjoys company.

Political: A person with a high political value finds his or her chief interest in controlling and organising people, and tends to enjoy running meetings and societies etc.

Religious: Finally a high religious value indicates a person who is relating his whole life to some central idea --- either a formal religious creed, or a philosophy of life.

THE LINKAGE COLOURS USED IN SECTION E
OF THE EXPERIMENT.

These squares of colour were used only on the agent's
Originals: they were not given to either group of subjects.

Monday,
September 15th:



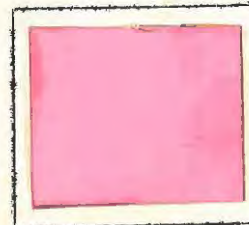
Tuesday,
September 16th:



Wednesday,
September 17th:



Thursday,
September 18th:



Friday,
September 19th:



(14) The specific instructions issued to the subjects
for each section of the experiment:

In addition to the general instructions that were issued unchanged each week, the subjects also received specific instructions on how each type of linkage material was to be used as it was introduced. It was thought advisable to tell them not only what they had to do, but also to explain something of the aim of each particular section, in the hope that this would help to maintain their interest in the experiment.

All subjects received exactly the same specific instructions, regardless of whether they had been allotted to the Control or the Experimental Group. The only difference was in the actual linkage material they received.

These specific instructions are given overleaf.

EXPERIMENTS IN PARANORMAL COGNITION: EXPERIMENT II.SPECIFIC INSTRUCTIONS FOR SECTION "A".FOR BOTH "E" AND "C" GROUP SUBJECTS.THE AGENT FOR YOUR GROUP:

For this investigation the subjects have been divided into two groups. Your group is asked to try and reproduce the drawings made by an agent in a house in KENILWORTH, CAPE TOWN. The members of the other group will be trying to reproduce the drawings of an agent in another part of the country.

Throughout the experiment think solely of your agent -- please don't attempt to find out details of the agent serving the other group.

AIM OF THE EXPERIMENT:

The investigation is designed to discover whether linking the subjects to the agents in various different ways has any effect on the number of the agent's drawings correctly reproduced by the subjects. It has been divided into five sections, for each of which different types of linkage will be provided.

AIM OF SECTION "A":

This first section is designed as a control, and is intended to find out how many drawings are correctly reproduced when the linkage between the subjects and their agent is very slight. The link between the agent and the subject in this first part of the experiment is the subject's knowledge of where the agent is operating.

WHEN TO START:

Your agent will make drawings for this section of the experiment in a house in KENILWORTH, CAPE TOWN, on:

Monday,	August 11th.
Tuesday	12th.
Wednesday	13th.
Thursday	14th., and
Friday	15th.

You should therefore make your first two drawings on Monday August 11th., and then two on each of the other days mentioned above --- ten in all for section "A".

SUBJECTS NOT IN RESIDENCE AT R.U.C.:

Are asked to bring or post their books, as each section is completed, to: Mr. M.C. Marsh, c/o Psychology Department, R.U.C., Grahamstown.

Tear off the bottom of this page here & pin it in your room:

HAVE YOU REMEMBERED YOUR

TELEPATHY
DRAWINGS

TODAY
?

EXPERIMENTS IN PARANORMAL COGNITION: EXPERIMENT II.SPECIFIC INSTRUCTIONS FOR SECTION "B".FOR BOTH "E" AND "C" GROUP SUBJECTS:WHAT YOU HAVE TO DO:

Each page in this book bears a different square of colour. As far as possible confine your attention to the page being used, and do not turn forward or back to look at the other colours.

Similar squares of colour will appear on the drawings your agent makes. Provided that you use the sheets in their correct numerical order, the square of colour appearing on your agent's drawings each day will be absolutely the same as that appearing on the sheet you use.

The other group of subjects will be using an entirely different set of colours. Please do not attempt to find out what their colours are: it is essential that you avoid seeing them.

In this section of the experiment your agent will associate each pair of target drawings made with the square of colour appearing on your counterfoil.

When trying to get impressions of what your agent has drawn, therefore, please fix your attention on the square of colour appearing on your counterfoil.

In all other respects make your drawings in the same way as in section "A".

Your agent will make drawings for this section of the experiment on :

Monday August 18th,
Tuesday August 19th,
Wednesday August 20th,
Thursday August 21st, and
Friday August 22nd.

You should therefore make your first two drawings on Monday August 18th, and then two on each of the other days mentioned above --- ten drawings in all for section "B".

WHY YOU HAVE TO DO THIS:

The experiment this week is designed to find out whether the association of ideas plays any part in telepathy. Normally we tend to link words together in our minds, like "Tom, Dick and Harry", or "Bread and Butter". If one of these associated words is brought to our attention, we tend to think automatically of the others.

Suppose now we printed the word "Butter" on this book, and on the sheet used by the agent for a drawing. When the drawings come to be made, it is likely that the agent will draw a loaf of bread, and that you will reproduce the same object, due to the word "Butter" calling up the same association "Bread" in both your minds. Here the word "Butter" is a linkage idea between you and your agent.

Thus only two requirements are necessary for both you and your agent to produce drawings of the same object:

These two requirements are:

- (1) There must be a linkage idea presented both to the agent and the subject, and
- (2) Previous experience should have associated the linkage idea with the same object in the agent's and subject's minds.

Now in a telepathy experiment such as this, the first of these requirements is fulfilled. The agent, when making a drawing, knows that a telepathy experiment is taking place, just as you too feel that you are taking part in one. Thus the linkage idea present to both the subject and agent is the "Idea of the Experiment".

Now previous experience has not connected this idea very strongly with any particular things or objects, either in your mind, or your agent's, so that the second requirement is not fulfilled.

However, as each drawing is made, it will become quite strongly associated in the agent's mind with the "Idea of the experiment", and also with any more specific objects present at the time the drawing is made. Consequently it only becomes necessary to make one assumption to apply the above explanation to the present experiment. This assumption is

THAT THE ASSOCIATIONS FORMED IN THE AGENT'S
MIND BECOME OPERATIVE IN THE SUBJECTS'
MINDS AS WELL.

It is this assumption that we are going to test in this section of the experiment.

In section "A" of the experiment, done last week, linkage was at a minimum. The ten objects drawn were associated only vaguely and generally in your agent's mind with the "Idea of the Experiment". This week the agents are going to form associations much more precisely. The drawing sheets they use will each bear a square of colour, and a different square of colour will be used daily. In this way your agent will associate each pair of drawings with a particular colour ---- the same colour appearing on your counterfoil each day.

If associations are important, as the association theory postulates, these more numerous and definite linking ideas should cause the results of section "B" to be better than those of section "A".

When you are deciding what to draw, therefore, please keep your attention fixed on the linkage idea provided --- the square of colour on your counterfoil.

EXPERIMENTS IN PARANORMAL COGNITION: EXPERIMENT II.SPECIFIC INSTRUCTIONS FOR SECTION "C"FOR BOTH "E" AND "C" GROUP SUBJECTS.AIM OF SECTION "C":

One of the most interesting branches of Psychical Research is Psychometry. In this type of experiment the subjects are supplied with a personal object belonging to their agent. The function of this personal object seems to be to put the subjects in contact with the mind of the person to whom the object belonged.

Anthropologists and students of magic will find the idea a familiar one. The way in which these objects function has also been studied experimentally by Dr E. Osty and also by Dr. Hettinger at London University. Dr. Hettinger published his findings in two books: "The Ultra-perceptive Faculty" and "Exploring the Ultra-perceptive Faculty"

This week we are attempting to find out whether providing you with an object which once belonged to your agent will improve your results or not.

WHAT YOU HAVE TO DO:

At the bottom of this sheet you will find a square of material from one of your agent's handkerchiefs.

While you are trying to get impressions of your agent's drawings in this section of the experiment, take this bit of material between your fingers, and keep in contact with it while you are making your drawings.

It is essential NOT to come in contact with the objects supplied to the other group of subjects, which link them to the other agent.

In all other respects make your drawings in the same way as you did in sections "A" and "B" of the experiment.

Your agent will make drawings for this section of the experiment on:

Monday August 25th.,
 Tuesday August 26th.,
 Wednesday August 27th.,
 Thursday August 28th., and
 Friday August 29th.

You should therefore make your first two drawings on Monday August 25th., and then two on each of the other days mentioned above ---- ten different drawings in all for section "C".

THIS SQUARE OF MATERIAL COMES FROM
 A HANDKERCHIEF ONCE OWNED AND USED
 BY THE AGENT SERVING GROUP " "
 SUBJECTS:

EXPERIMENTS IN PARANORMAL COGNITION: EXPERIMENT II.SPECIFIC INSTRUCTIONS FOR SECTION "D"FOR BOTH "E" AND "C" GROUP SUBJECTS.WHAT YOU HAVE TO DO:

A photograph and a description of your agent will be found in the booklet provided.

Before making your first two drawings read through this information given by your agent in the booklet.

It is essential that you should see only the information supplied on the agent serving your group. Do not attempt to find out what the other group's agent is like.

When you come to make your drawings each day, open the booklet at your agent's photograph, and have it before you while you draw. Notice particularly the face, and whether the box used to enclose the drawings after they have been made is marked in any way or not.

If you can spare the time skim rapidly over the salient points of your agent's personality each day before you make your drawings.

In all other respects make your drawings in the same way as you did for sections "A", "B", and "C" of the experiment.

Your agent will make drawings for this section of the experiment on:

Monday September 1st.,
Tuesday September 2nd.,
Wednesday September 3rd.,
Thursday September 4th., and
Friday September 5th.

You should therefore make your first two drawings on Monday September 1st., and then two on each of the other days mentioned above --- ten different drawings in all for section "D".

THE AIM OF SECTION "D":

It has been suggested that telepathy seems to take place most frequently between people who have some personal or emotional bond in common, such as close friends, twins, husband and wife, parents and children, and siblings.

This week the effect of providing such personal acquaintanceship links between the agents and their subjects is to be studied.

In case you have lost your original reminder slip, a second one is printed below which you can use.

TELEPATHY
DRAWINGS

EXPERIMENTS IN PARANORMAL COGNITION
EXPERIMENT II.

SPECIFIC INSTRUCTIONS FOR SECTION

"E"

FOR BOTH "E" AND "C" GROUP SUBJECTS.

THE AIM OF THIS FINAL SECTION:

In this section the aim is to find out whether the way in which the agent makes his drawings has any effect on the subjects' results. Consequently the linkage between the subject and the agent remains the same as that used for section "D". You will find that the same information and picture of your agent that you received for that section has been returned to you. Use this material in the same way as you did before.

As this is the final section of the experiment, it would be greatly appreciated if you would give certain additional information on your attitude towards the experiment and how you carried it out. Please, therefore, fill in the questionnaire supplied with these instructions, and return it with your drawings.

WHAT YOU HAVE TO DO:

A photograph and a description of your agent will be found in the booklet provided.

Before making your first two drawings read through the information given in the booklet.

It is essential that you should see only the information supplied to you on the agent serving your group. Do not attempt to find out what the other group's agent is like.

When you come to make your drawings each day, open the booklet at your agent's photograph, and have it before you while you draw. Notice particularly the face, and whether the box used to enclose the drawings after they have been made is marked in any way or not.

If you can spare the time skim rapidly over the salient points of your agent's personality each day before you make your drawings. In all other respects make your drawings in the same way as you did in sections "A" "B" "C" & "D" of the experiment.

After making your last two drawings, fill in the questionnaire which has been supplied to you.

Your agent will make drawings for this section of the experiment on:

Monday September 15th,
Tuesday September 16th,
Wednesday September 17th,
Thursday September 18th, and
Friday September 19th.

You should therefore make your first two drawings on Monday September 15th, and then two on each of the other days mentioned above ---- ten different drawings in all for section "E".

(15) Diaries kept by the subjects, experimenter and agent:

The subjects were required to give information daily on their counterfoils as to their physical condition, the amount of drugs they had consumed such as alcohol, coffee, tea, etc., during the course of the day, and were also asked to report on the degree of confidence they felt in their drawing, and whether they had experienced a clear image in their mind's eye or not. They were also invited to add any other comments that they felt were relevant.

The experimenter also kept a diary of the daily happenings in Grahamstown while the experiment was in progress, in which he noted anything that might tend to increase or decrease the frequency of any of the possible originals. This diary was kept informally, but precautions were taken to see that it was always locked up. The experimenter had no knowledge, while the experiment was in progress, of which of the 100 originals were being selected by the agent for use as targets, and so could not have caused any sensory leakage to the subjects through normal sense channels.

The agent was also required to keep a diary of his activities at the transmitting end. In order to make sure that all the relevant points had been considered by him each day, a printed form was supplied to him, together with instructions on how to fill it in. Copies of these documents are given on the next two pages.

EXPERIMENTS IN PARANORMAL COGNITION.Instructions to the Agent on how to fill in the
Diary.

Please make an entry for each day you make drawings.

Fill in part 1 of the diary at the time you make the drawings, and part 2 just before going to bed of an evening.

In part 1 you are asked to give the serial numbers of the drawings, their titles, the time they were made, and the date on which they were made. You are also asked to indicate whether you found the objects easy or difficult to draw, and whether they called up any associations in your mind while you were making them.

In part 2 you are asked to give the time you are making the entry, and then whether you thought of the objects at all during the day, and if so, when. There is a possibility that your physical condition might affect your sending efficiency, so you are asked to indicate whether you feel normal, ill in any way, tired, or in pain.

Similarly, psychological factors may have an effect, so you are asked to indicate whether you experienced any outstanding psychological feelings, such as worry, excitement, emotions during the course of the day.

Previous investigators have suggested that various drugs may also have an effect on your sending capacity, so please indicate your consumption of them during the day.

I would like you as far as possible to maintain the same routine and conditions throughout the whole experiment, but if you should unwittingly vary the procedure, or relax the precautions, indicate that you have done so in the diary.

Finally you are asked to make any other comments or notes you think may be relevant; the reverse of the page may be used for this too if necessary.

AGENT'S DIARY.

PART I: (To be filled in at the time of making the drawings).

Serial No. of first drawing: Title of first drawing:

Serial No. of second drawing: Title of second drawing:

Today's date: Time of making drawings:

Ease or difficulty of making drawings:

Any associations called up by the drawings:

PART II: (To be filled in just before going to bed).

What time are you making this entry ?

Did you think of the drawings during the course of the day ? If so, specify when:

What was your physical condition today ?

Your mental condition ?

Consumption of drugs (from making drawings to time of writing):

Cigarettes: Cups of tea: Gups of coffee:

Alcohol (specify):

Other drugs (Asprin etc.) (specify):

Did you make any changes in the routine procedure & precautions today ? If so, specify:

Add any other information you would like to give on the reverse of this sheet.

HOW THE DATA WAS SCORED:

- (1) Preparation of the subjects' drawings for assessment.80
- (2) Preparation of the originals for use by the judges.80
- (3) The assessment of the subjects' drawings by the three primary judges.90
- (4) Combination of the assessments of the three primary judges to give one score.92
- (5) Mrs Forster's check on the Type A hits awarded by the three primary judges.92
- (6) Dr. Humphrey's and Mrs. Forster's check on the possibility of the judges being able to distinguish Control from Experimental Series Originals on the basis of artistic style.92
- (7) The use of Control Series Originals as a safeguard.93

(1) Preparation of the subjects' drawings for assessment:

In order to eliminate bias on the part of the judges, it was necessary to make sure that they had no way of knowing whether any particular drawing was made by an experimental or a control group subject, or during which section of the experiment it had been made. This was accomplished as follows. The drawings were first of all detached from their counterfoils. Next the top of the sheet was folded over four times, concealing the true serial number of the drawing, and a piece of gummed paper was then stuck over the fold to keep it down, and also to make it impossible to see what the serial number was if it was accidentally or intentionally held up to the light.

The drawings were then randomised by being sorted into alphabetical order according to title. Finally each drawing was stamped with a second serial number in order from A to Z, so that each drawing could be identified during the assessment process and the assessments recorded easily.

Since the experimenter intended to be one of the three judges, this work was undertaken by a clerical assistant, and the experimenter did not see any of the subjects' drawings until they had been properly randomised.

(2) Preparation of the Originals for use by the Judges:

By the end of the experiment, the agent had drawn 50 Experimental Series Originals, which were used as targets in the experiment. This left 50 title slips still in the selector box. The agent then forwarded his originals, together with all the other experimental material he had used, including the selector box, direct to Miss Y. Fleischer. Miss Fleischer shook the remaining 50 titles from the box, made outline drawings as far as possible in the same style as those made by the agent of these objects, and then made tracings of the drawings the agent had made.

By making these tracings it was hoped that idiosyncrasies of the agent's style would be covered up, making it more difficult for

any person to tell drawings made by Miss Fleischer from those made by the agent. There were now 100 drawings supplied by Miss Fleischer: 50 Control Series Originals which she had drawn personally, and 50 Experimental Series Originals which she had traced. The drawings belonging to the Control and Experimental Series were now intermixed by being sorted into alphabetical order according to their titles, and were mounted in this randomised order on stiff cards, and sent to the judges.

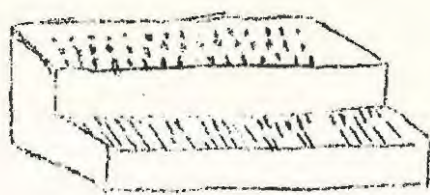
The writer has often been asked why he did not get one person to make both sets of drawings; it has been suggested that the agent could very well have made the remaining 50 Control Series Originals after the experiment had been completed. The writer was afraid that if he used the same person to make both sets of drawings the subjects might tend to score precognitive hits on the Control Series Originals as well as on the Experimental Series Originals at which they were really aiming. Since there was a possibility that the subjects' hits would be widely displaced in time, the experimenter wanted to keep the two sets of originals as distinct psychologically as possible, and thus in two different minds, rather than in only one.

In actual fact the E.S.P. experimenter, when exposed to the complications of precognitive telepathy and clairvoyance is in an extremely difficult position. Even after the experiment is over, any person who learns which were experimental and which were control originals may possibly provide an uncontrolled source from which the subjects may have drawn precognitive knowledge during the course of the experiment. It could be argued by a critic accepting the possibility of precognition, that the subjects gained their knowledge not from the agent, but from Miss Fleischer, the experimenter, and indeed all the people who read this report. Such possible sources of precognitive knowledge become increasingly difficult to control if the results of any experiment are communicated fully

to other workers in the field. Happily in the case of this experiment, the results themselves show that the linkage factors provided between subjects and agent were effective in increasing their scoring rate, proving that the subjects were getting their impressions from the mind of the agent, and not from other sources.

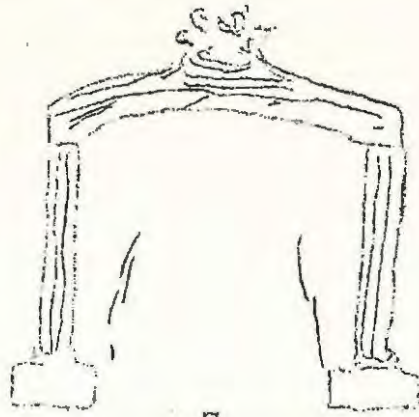
The 100 drawings, as drawn or traced by Miss Fleischer and supplied to the judges are given on the next seven pages. They have been photographically reduced to a little over half size. Readers are invited at this stage to try and separate them into two groups, on the basis of artistic style, etc., the one group being Control Series Originals, drawn by Miss Fleischer, the other Experimental Series Originals, drawn by the agent during the experiment. One source of criticism raised against the experiment has been that it was possible for the judges to make this distinction while assessing the subjects' drawings, and thus be more lenient in allowing hits on the Experimental Series Originals than on the Control Series Originals. By attempting to make such a separation at this stage readers can themselves arrive at an estimate of how applicable this criticism is. A list of Control and Experimental Series is given later, on page 330.

Dr. Betty Humphrey and Mrs. Forster also kindly undertook to check on this possible source of criticism, and the results of their separation are given later on page 121.

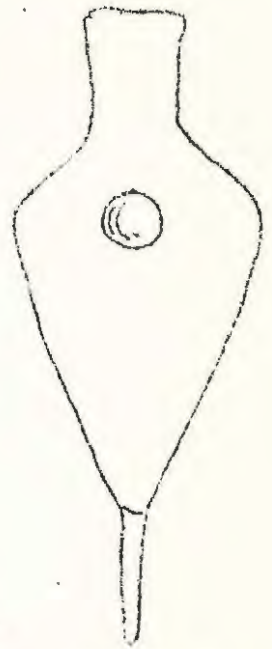


1 ACCORDEON

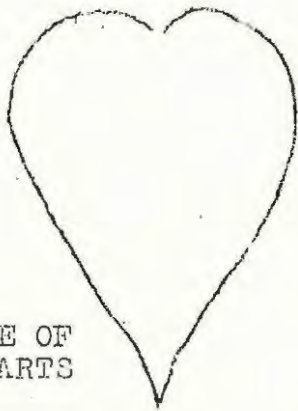
A



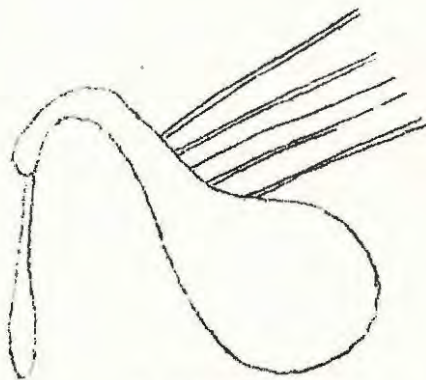
7
ARC DE TRIOMPHE



12 BELLOWS



2 ACE OF HEARTS



8 BAG-PIPES

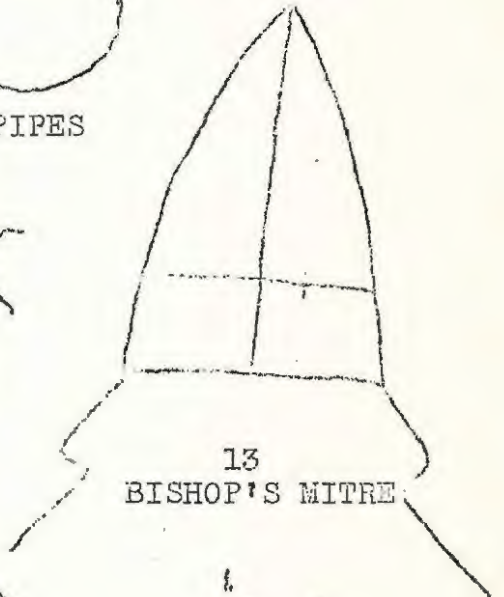
A



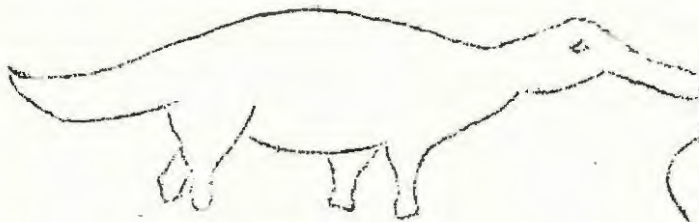
3 ALADDIN'S LAMP



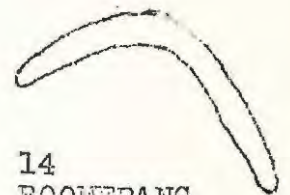
9 BARBED WIRE BARB



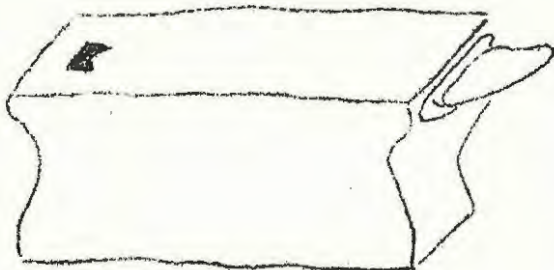
13
BISHOP'S MITRE



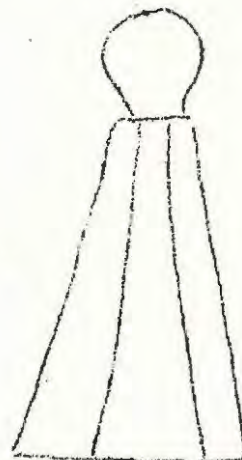
4 ANT BEAR



14
BOOMERANG



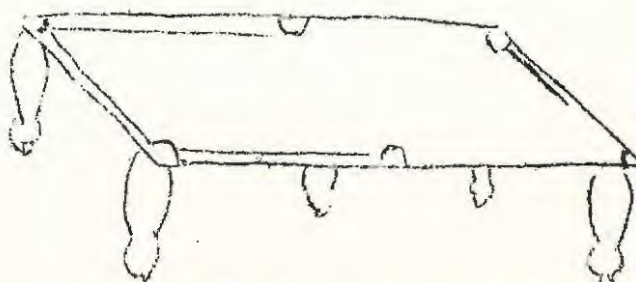
5 ANVIL



10 BEACON



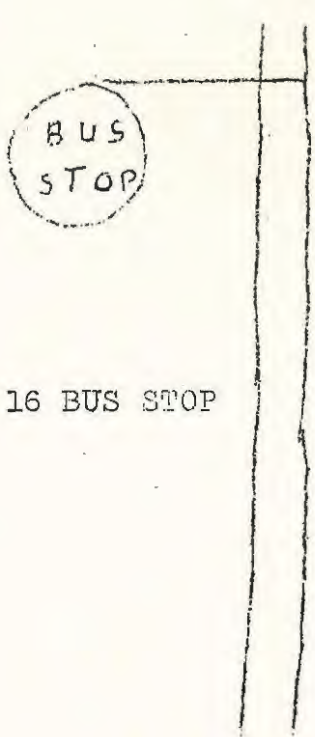
6 ASTERISK



11 BILLIARD TABLE



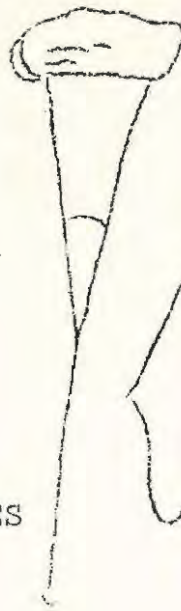
15
BOTTLE
OPENER



16 BUS STOP

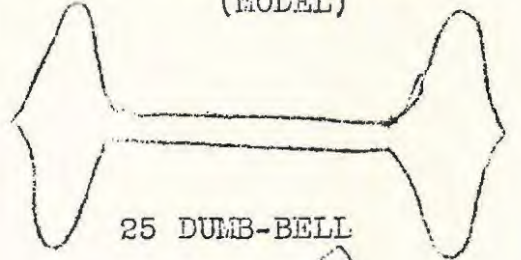


21 CRUTCHES



84.

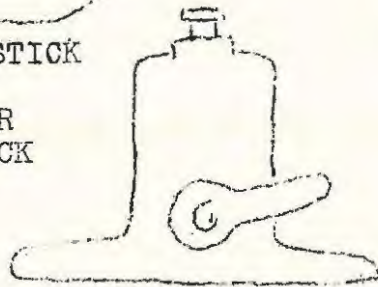
24 DISCUS THROWER (MODEL)



25 DUMB-BELL



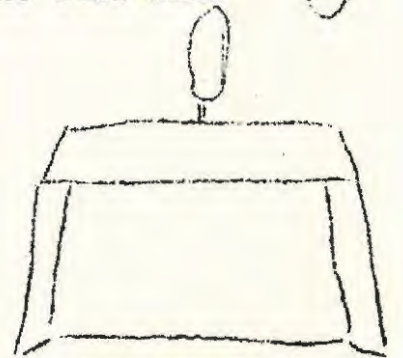
17 CANDLE STICK



18 CAR JACK



22 CUPID

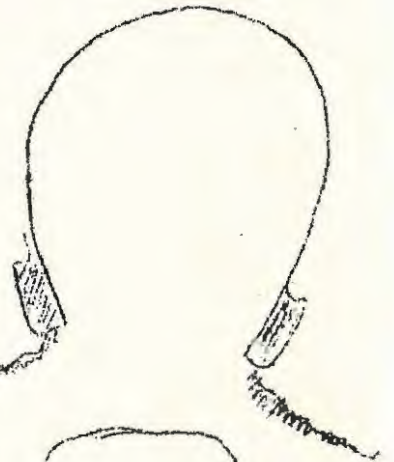


26 DUST PAN

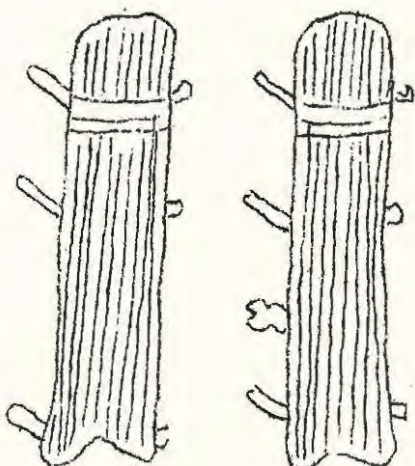
19 MR. CHAD



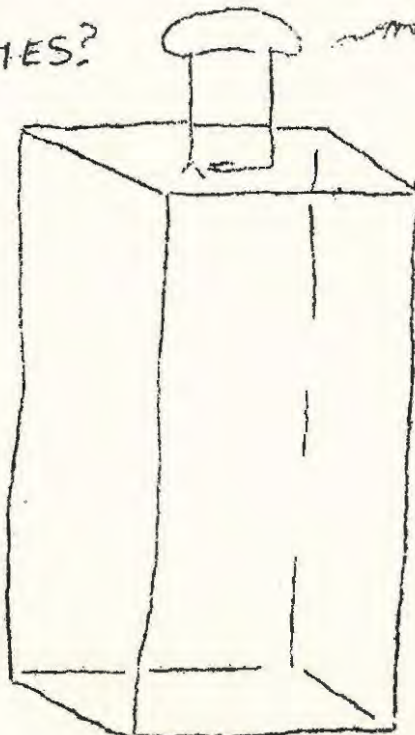
27 EARPHONES



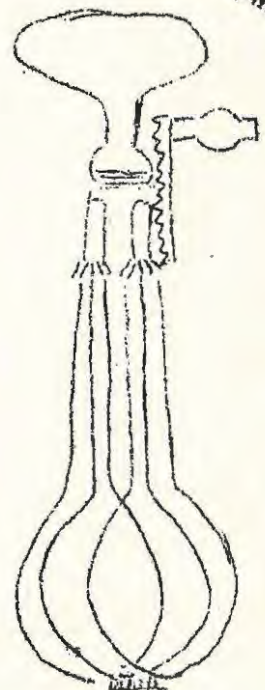
WOT! NO NAMES?



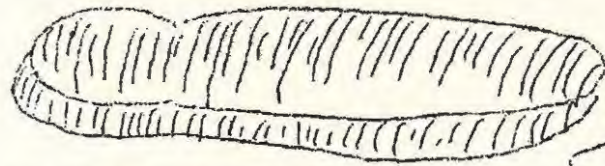
20 CRICKET PADS



23 DECANTER

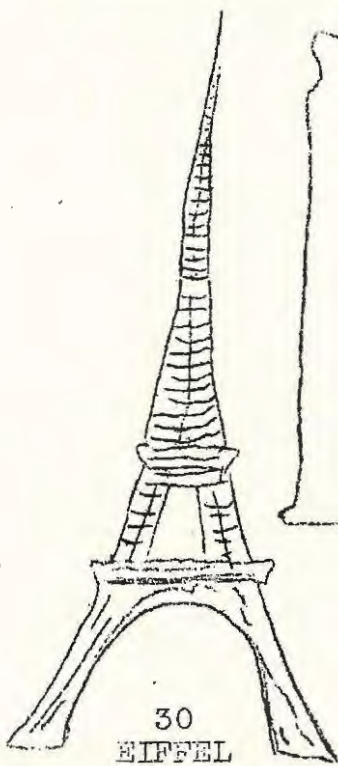
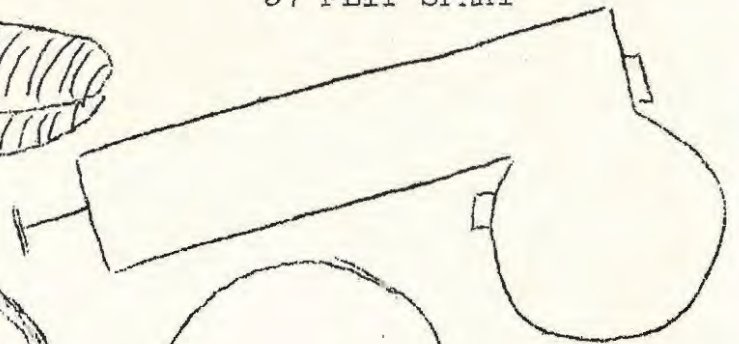


28 EGG-BEATER.

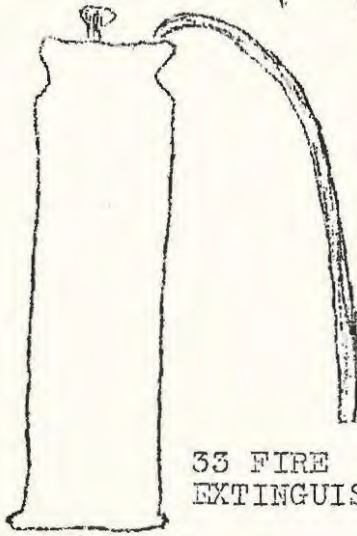


29 EGYPTIAN MUMMY

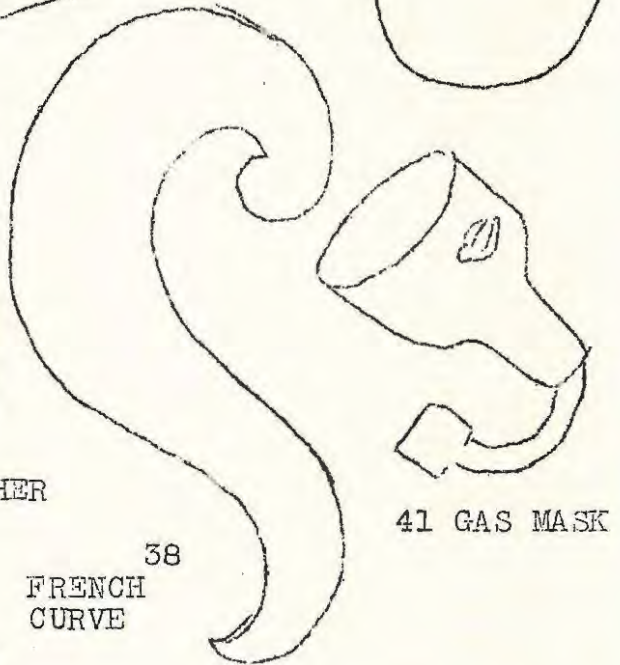
37 FLIT SPRAY



30 EIFFEL TOWER



33 FIRE EXTINGUISHER

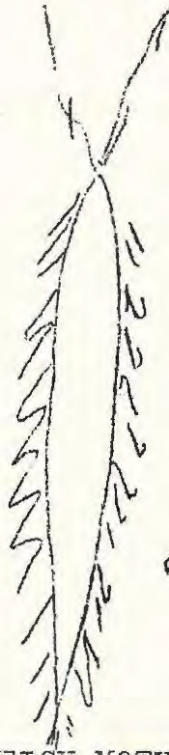


41 GAS MASK

38 FRENCH CURVE



31 FATHER CHRISTMAS



34 FISH MOTH



39 GOD PAN



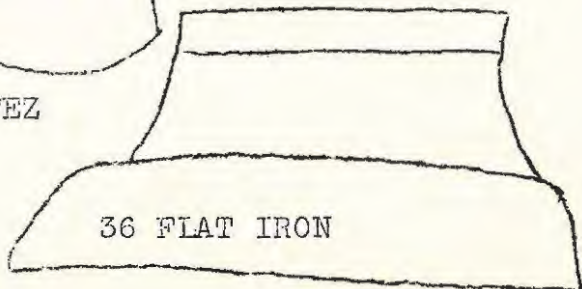
32 FEZ



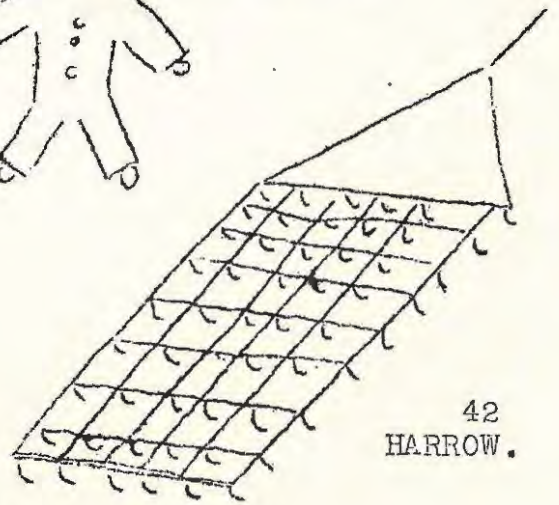
35 FISHING GAFF



40 GOLLYWOG

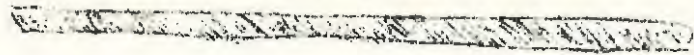


36 FLAT IRON

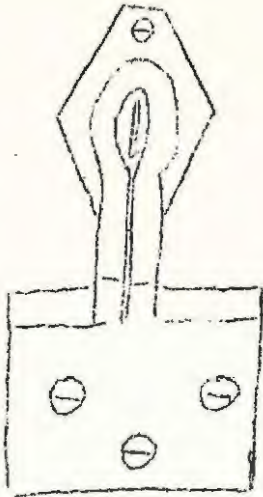
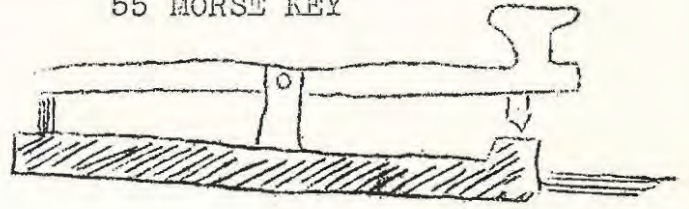


42 HARROW.

48 JAVELIN



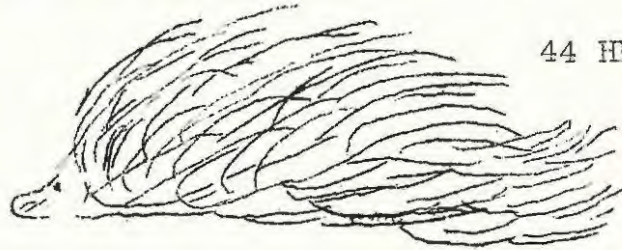
55 MORSE KEY



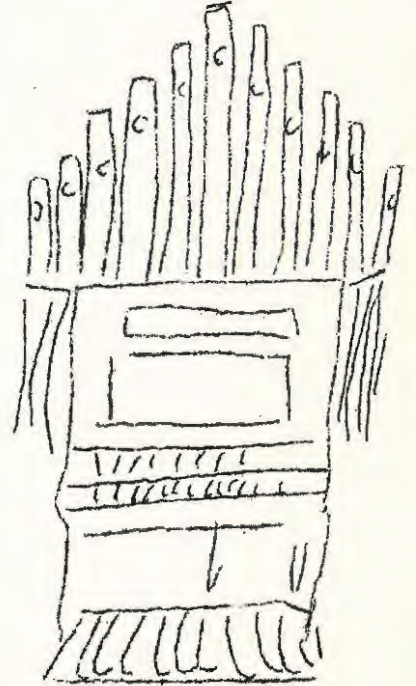
45 HASP AND STAPLE



49 JODPHURS

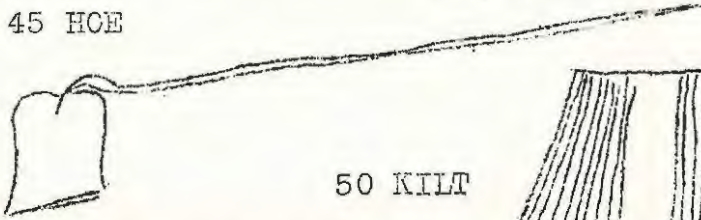


44 HEDGEHOG

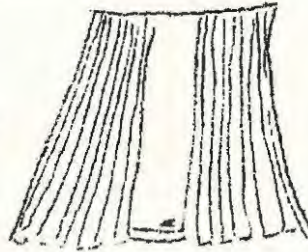


56 ORGAN

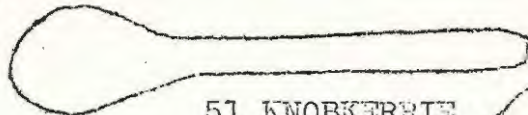
45 HOE



50 KILT



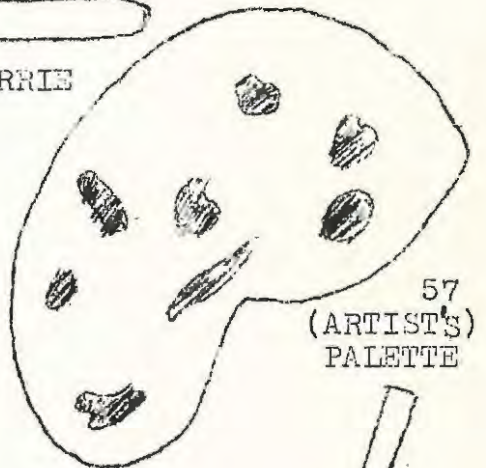
46 HONEYCOMB



51 KNOBKERRIE



52 LAUREL CROWN



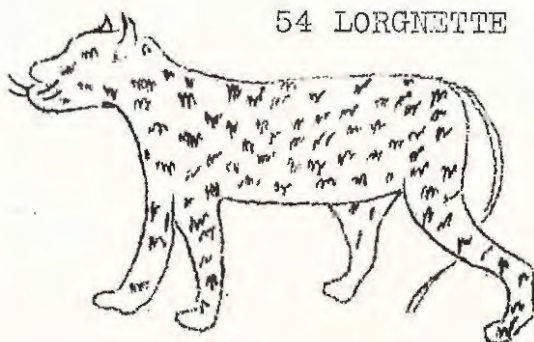
57 (ARTIST'S) PALETTE



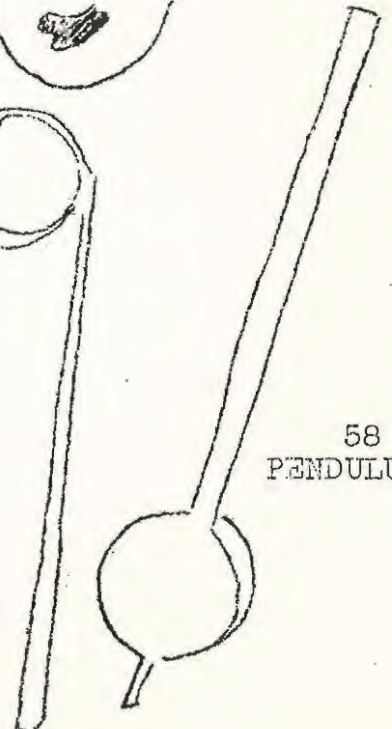
47 HOUR GLASS



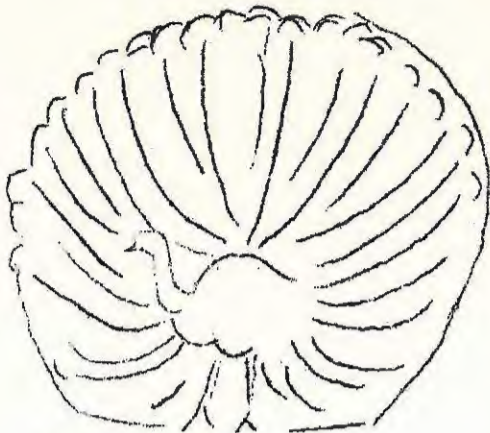
54 LORGNETTE



53 LEOPARD



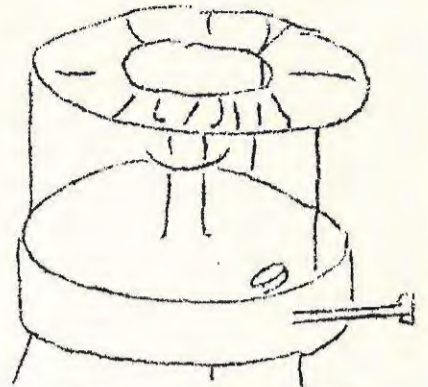
58 PENDULUM



59 PEACOCK

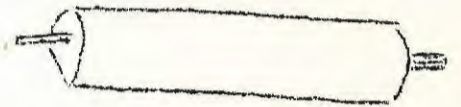


63 PRICKLY PEAR

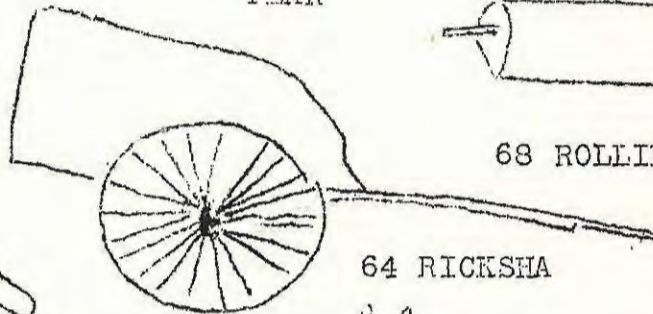


67 PRIMUS STOVE

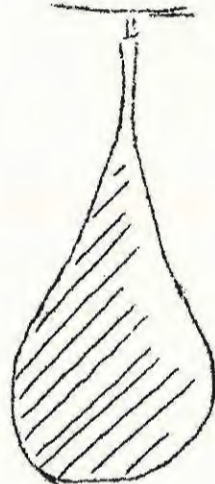
60 PENGUIN



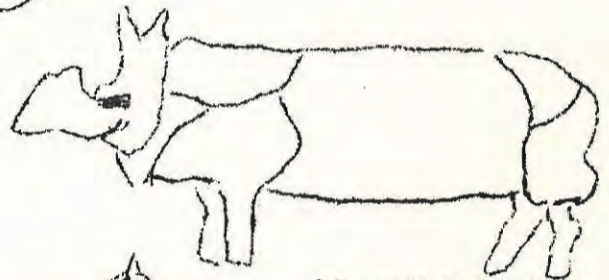
68 ROLLING-PIN



64 RICKSHA



65 PUNCH BAG



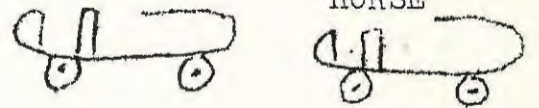
69 RHINOCEROS



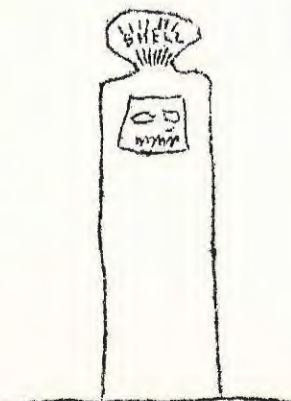
61 PHOENIX



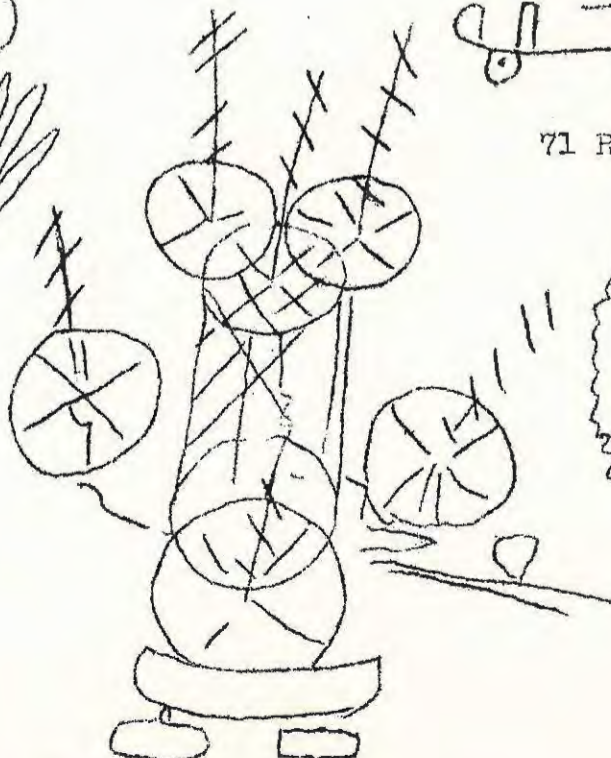
70 ROCKING HORSE



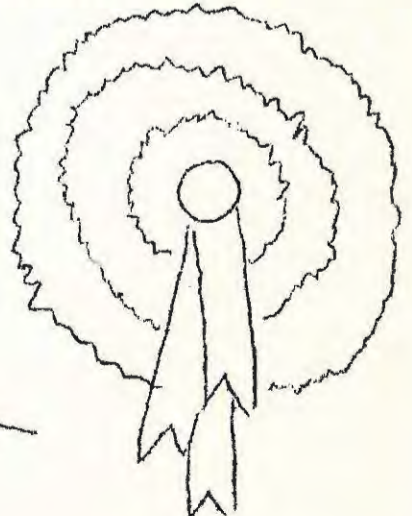
71 ROLLER SKATES



62 PETROL PUMP



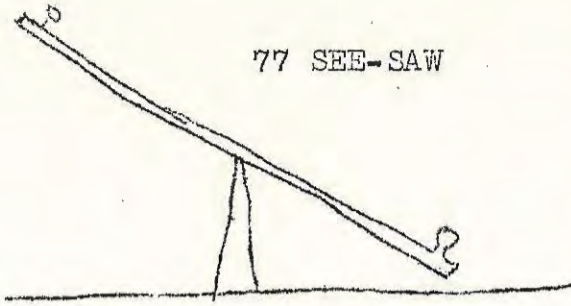
66 RADAR AERIAL



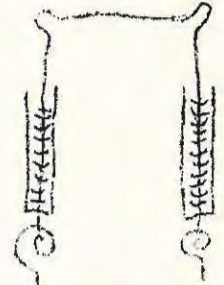
72 ROSETTE



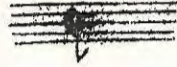
73 SAINT



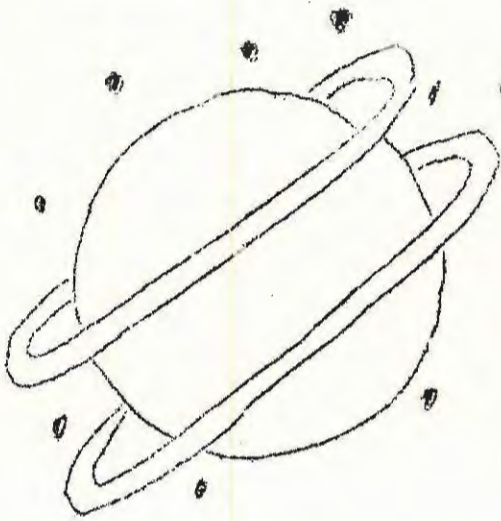
77 SEE-SAW



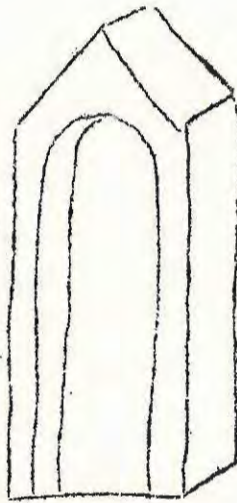
82 SPIFFY



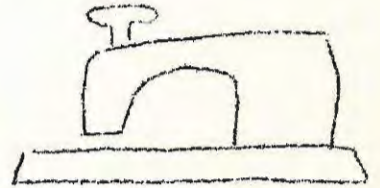
78 SEMI-QUAVER



74 SATURN



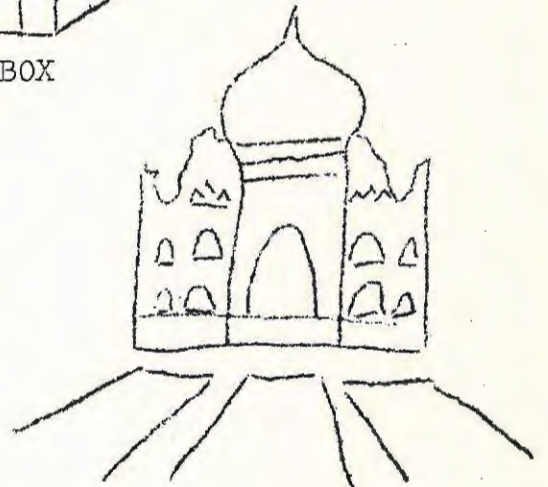
79 SENTRY BOX



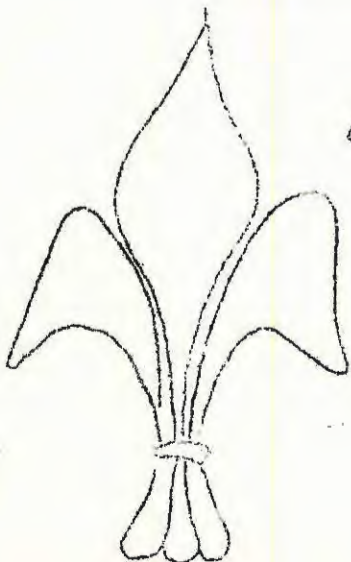
83 STAPLING MACHINE



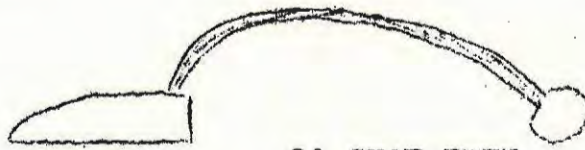
75 SCORPION



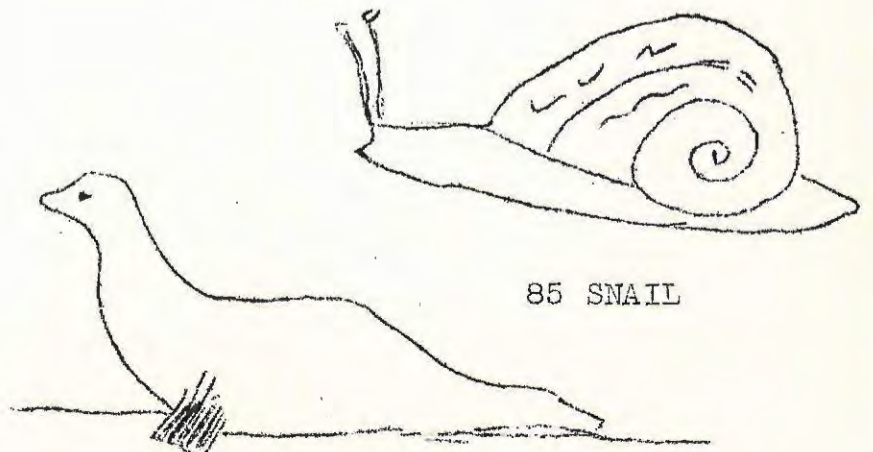
84 TAJ MAHAL



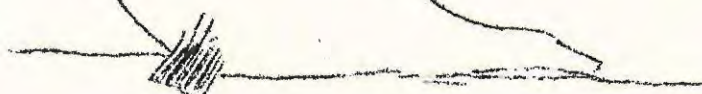
76 SCOUT BADGE



80 SHOE-TREE



85 SNAIL



81 SEAL



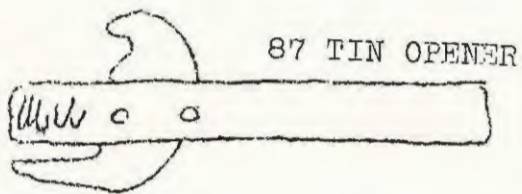
86 SPHINX



91 U-TUBE



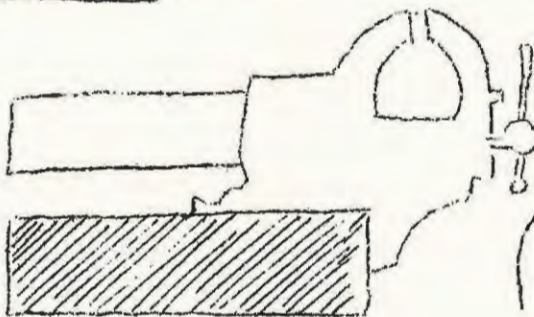
89.
97 WATER BAG



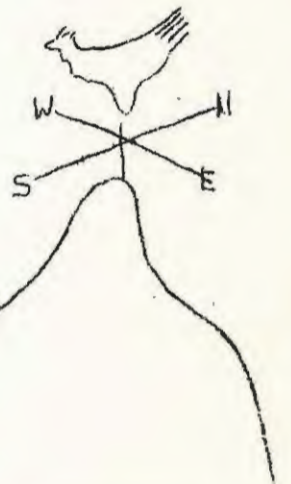
87 TIN OPENER



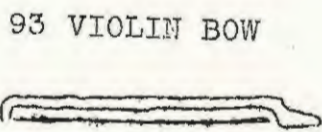
88 TRIDENT



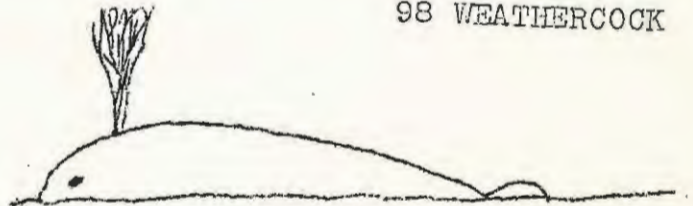
92 VICE



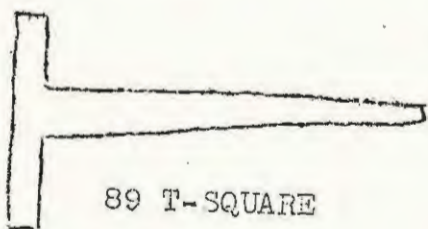
98 WEATHERCOCK



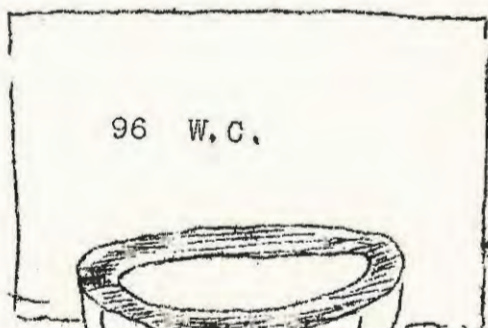
93 VIOLIN BOW



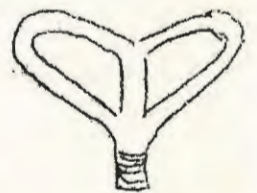
95 WHALE



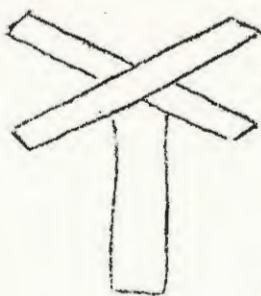
89 T-SQUARE



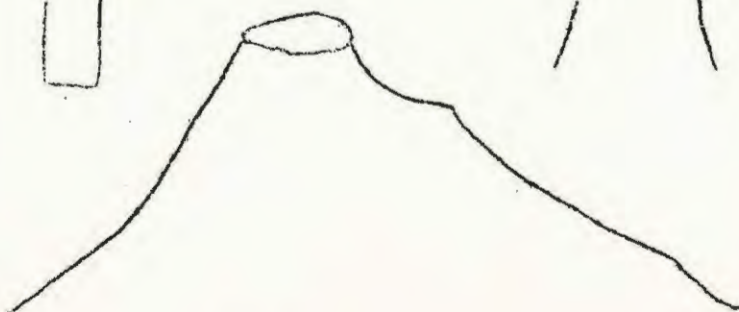
96 W.C.



99 WING NUT



90 TURNSTILE



94 VOLCANO



100 WISH-BONE

(3) Assessment of the subjects' drawings by the three primary judges:

Both the subjects' drawings and the originals were now ready for the judges' assessment. The experimental procedure had been so arranged that the experimenter himself could undertake the assessment of the drawings, but he felt that as an additional safeguard it would be advisable to have the material assessed independently by two other judges as well. This very considerable labour was kindly undertaken by the writer's mother, Mrs. L. F. Marsh, and his brother, Mr. D.R.C. Marsh.

The three judges compared each of the 17,440 subjects' drawings with each of the 100 originals, and classified any resemblances according to the scheme outlined on the next page. Each judge worked independently through all the subjects' drawings, and made the assessments without consulting either of the other two judges.

The procedure that the judges were to follow was carefully defined in a sheet of instructions. These instructions were based on those issued by Whately Carington to his judge, Mr. Hindson (13), but the actual types of hit looked for by the judges were wider than the standards adopted by Carington. This modification had been suggested as desirable by the work of Lettinger and Warcollier, which has already been summarised elsewhere. Briefly it will be recalled that their work suggested that the telepathic message was subject to considerable distortion before it became conscious. Hence it was thought that it would be of value to include categories of hit which would cover these fainter resemblances.

The full instructions given to the judges are set out on the next page.

(4) Combination of the assessments of the three primary Judges into one score:

After all three judges had completed their assessments, the judgements of all three were combined, by accepting the majority decision in disagreement, or by accepting the lowest of the assessments if all three judges disagreed on the type of hit which should be allowed on a particular drawing. This happened very seldom. Examples:

AF3	AF3	AF3	Assessment adopted:	AF3.	
AF3	AF2	AF2	"	"	AF2
AF3	AF2	AF1	"	"	AF1

In cases where there was total disagreement on any type of title or association hits, sometimes an arbitrary decision had to be made as to the most justifiable of the three assessments. If all three seemed equally justified, the decision of the first judge was accepted arbitrarily. Total disagreement of this sort very seldom occurred, however, so that the necessity for this procedure very seldom arose. *

(5) Check on the A Hit assessments awarded by the three primary Judges:

As an additional safeguard, Mrs. Forster kindly undertook to review all the A hits awarded by the judges, and to indicate any which in her opinion were not sufficiently close to the target title to be classified as A hits. This review will be given fully in the "First Main Analysis" which follows shortly.

(6) Check on the possibility that the judges could tell that some of the originals had been made by one person, and some by another:

Dr. Betty Humphrey and Mrs. Forster kindly undertook to test this possibility for the writer, and the results are also given in the "First Main Analysis".

* Reproductions of a sample of subjects' drawings and the hits allocated will be found in the Fifth Main Analysis.

(7) Why a Control Series of Originals was used as well as the Experimental Series of Originals.

In outlining this experimental procedure, we have already discussed the use of the Control Group of Subjects. It remains to explain what function the introduction of the Control Series of Originals was intended to perform.

It seemed to the writer that although the Preliminary Frequency Experiment had determined the chance expectancy of drawing any of the originals, the bare use of this previously determined chance probability in determining the significance of the results might be open to valid criticism. If it were found that the subjects scored more hits than would be expected on this previously determined basis, it might be argued that this increased score was not due to E.S.P., but to other factors, such as seasonal changes occurring between the administration of the Preliminary Frequency Experiment, and the Main Experiment, or chance uncontrolled factors such as a fire, which would make the subjects think of fire engines and fire extinguishers more often than they had done during the Preliminary Frequency Experiment.

To find out whether any such chance factors were influencing the results, it was decided to introduce the Control Series of 50 originals drawn by Miss Fleischer, and score all the subjects' drawings not only against the 50 Experimental Series Originals, but against the Control Series as well. It was reasoned that any chance factors affecting the frequency of hits on the Experimental Series Originals, considered as a group, would also affect the frequency of hits on the Control Series Originals, considered as a group, in the same way.

Thus if both the Experimental Series Originals and the Control Series Originals had significantly more hits scored on them than could be expected from the previously determined chance probability, there would be no reason for believing that the hits scored on the Experimental Series were produced by

anything more than chance factors. If, however, the subjects scored significantly more hits on the Experimental Series Originals than the chance frequency determined in the Preliminary Frequency Experiment, whereas the hits scored on the Control Series Originals did not differ significantly from the previously determined chance frequency, then we could be sure that the hits scored on the Experimental Series Originals were not merely caused by uncontrolled chance factors operating. Unless it could be shown that some uncontrolled chance factors were operating selectively on the Experimental Series Originals, and not on the Control Series Originals, we would be justified in attributing the surplus of hits scored on the Experimental Series Originals as due to the subjects exercising E.S.P.

Thus a comparison of hits on these two series of originals will give some indication of the influence of any chance factors. As will be seen later in the analysis of the data, the effects of such chance factors show up in other ways as well, and can be allowed for in assessing the results.

COMPARISON OF THE FREQUENCY OF HITS SCORED ON
EXPERIMENTAL AND CONTROL SERIES ORIGINALS.

(1) <u>Introduction:</u>	98
(2) <u>Significance of the Raw Score Differences:</u>	
(a) Explanation of the statistical method used:.....	99
(b) Calculations C1 to C20:.....	105
Table I:.....	107
(c) Conclusions drawn from the calculations:.....	108
(3) <u>Factors which might account for the observed differences:</u>	
(a) Relating to the way in which the originals were selected from the data of the Preliminary Frequency Experiment:.....	108
(b) Relating to actions on the part of the agent:.....	112
(i) Deliberate conscious selection of certain originals and rejection of others:.....	112
(ii) Unconscious use of P.K. and Precognitive E.S.P. to select favourable originals:.....	113
(c) Relating to the subjects:	
(i) The "Bunching Effect" and corrections:.....	115
(ii) Rational inference and sensory leakage:.....	119
(iii) The tendency for subjects to score A hits on the originals they had suggested in the Preliminary Frequency Experiment:.....	120
(d) Relating to actions on the part of the judges:.....	121
(i) Possible bias arising from ability to tell which were Experimental and which Control Series Originals:.....	121
(a) By inspection of Originals: based on differences of artistic style of the two artists responsible:.....	121
(b) By the use of Psi ability:.....	123

FIRST MAIN ANALYSIS
(continued)

(3) <u>Factors which might account for the the observed differences (continued):</u>	
(d) Relating to actions on the part of the judges (continued):	
(ii) By the judges adopting too lenient a standard in assessing A hits:	124
(a) Mrs. Forster's review of the A hit awards:	124
(b) Corrections based on her recommendations:	127
(e) Summary of all the corrections made:	128
(4) <u>Review of further methods of assessment to be employed:</u>	129
(5) <u>Calculation of the significance of the difference in number of fully corrected A hits scored on Experimental as compared with Control Series Originals:</u>	131
(a) Introduction:	131
(b) Statistical Methods:	132
(c) Calculations C21A and C21B:	134
(d) Conclusions drawn from these calculations:	136
(6) <u>Calculation of the significance of the difference in numbers of Raw A hits expected in terms of the Preliminary Frequency Experiment, and those actually obtained on Control and Experimental Series Originals in the Main Experiment:</u>	137
(a) Introduction	137
(b) Statistical method:	138
(c) Calculations C22A and C22B:	139
(d) Conclusions drawn from the above calculations:	140
(7) <u>Calculation of the significance of the difference in numbers of Fully Corrected A hits expected in terms of the Preliminary Frequency Experiment, and those actually obtained on Control and Experimental Series Originals in the Main Experiment:</u>	140
(a) Introduction:	140
(b) Statistical Method:	140
(c) Calculations C23A and C23B:	141
(d) Conclusions drawn from the above calculations:	143

FIRST MAIN ANALYSIS:
(continued):

(8) <u>Assessment of the behaviour of the Control and Experimental Series Originals against the behaviour of a large population of similar series of Originals in the Main Experiment:</u>	146
(a) Introduction	146
(b) Statistical Method:	146
(c) Calculation C24:	148
(d) Conclusions drawn from Calculation C24:	149
(e) Table II:	151
(9) <u>Summary of the results of the First Main Analysis, and the conclusions drawn from the evidence:</u>	152
(a) Review of the assessments and the results:	152
(b) Summary of the evidence unfavourable to the telepathy hypothesis:	155
(c) Summary of the evidence supporting the telepathy hypothesis:	156
(d) Conclusions:	156

(1) Introduction:

Before the results are examined to discover what influence the linkage objects had on the scoring level of the subjects, or how this scoring level was related to personality factors, it is first necessary to find whether or not the results show any evidence that E.S.P. ability was being used by the subjects. The experiment was planned in such a way as to make possible a number of independent methods of testing the results to decide this question .

It is perhaps logical to start with an examination and comparison of the number of hits scored on the Control Originals (which were not targets in the experiment) with those scored on the Experimental Series Originals actually drawn by the agent. Unfortunately this is the most complex and involved of all the analyses that will be undertaken, necessitating several corrections of the raw scores. In order to appreciate the nature of this analysis it will be necessary to recapitulate briefly parts of the experimental procedure.

It will be remembered that from the Preliminary Frequency Experiment lists, 93 objects were selected that had occurred only once in the 15,835 titles, and 7 which had occurred only twice. From this group of possible target objects 50 were selected by an essentially random method to be target or "experimental" originals, and were drawn by the agent during the course of the experiment. The remaining 50 titles were used as control originals, and drawings were only made of them after the experiment had been completed. Before being sent to the two judges the two sets of originals were thoroughly intermixed, so that the judges, when they came to allocate their hits, would not know whether they were allocating hits on Control or Experimental Series Originals, and thus could not be biased. The judges then compared all the subjects' drawings, with all 100 of these originals, awarding various types of hit according to the degree of resemblance.

Now assuming that the subjects were equally likely to make drawings on a chance basis from the Experimental and Control Series Originals, and assuming that the procedure adopted successfully excluded bias on the part of the judges, then one would expect that any E.S.P. ability exercised by the subjects would cause them to score hits on the Experimental Series Originals, which were their targets, rather than on the Control Series Originals, which had not been used in the experiment at all.

Consequently it will first be necessary to examine the results to find whether there are any significant differences between the number of hits scored on the two series of originals. If such differences are found, it will then be necessary to examine the assumptions that have been made in detail, in order to try and lay bare the factors responsible for the differences.

This preliminary analysis will also show what types of hit are most likely to bring out any E.S.P. effects in the material. Whately Carington (13) found that it was better to adopt a fairly strict standard in awarding hits, and disregard the fainter resemblances, but there is no way of knowing at this stage whether this standard will also give the best results in this investigation. One would expect that the strict judgement categories like F3, G3, H3 and especially A would give the best results, and that the fainter resemblances like F 1 & 2, G 1 & 2, H 1 & 2, B, C, D, E, and I, J and K would be less efficient, and consequently give less clear-cut results. For this reason the different types of hit have been considered separately in this analysis.

(2) Significance of the differences in Raw Scores:

(a) Explanation of the statistical method used.

In the calculations that follow the Chi Square Statistic has been used to test the assumption that there is no significant difference in the number of hits scored on the Experimental Series Originals, and those scored on the Control

Series Originals, that, in effect, they come from the same statistical population.

As an example of the method, let us explain calculation C1 in detail. In the Main Experiment the judges awarded 133 A hits on Experimental Series Originals, and 37 on Control Series Originals, for the 17,440 trials made by the subjects. We wish to find whether this observed difference in scoring rate on the two series of originals is large enough to be regarded as significant.

To do this, it is necessary to find what scores would be expected if no difference existed (the "theoretically expected scores" in terms of the null hypothesis that no difference exists). We next compare our observed scores with these theoretically expected scores to find whether they differ significantly from them. We assess the significance of the difference by means of the Chi Square procedure, which involves the following steps:

- (1) Finding the theoretically expected score "E"
- (2) Finding the deviation "D" of the Observed score "O" from the theoretically expected score "E"
i.e., $D = O - E$
- (3) Squaring the deviation, D, to give " D^2 "
- (4) Dividing the squared deviation by the theoretically expected score thus: $\frac{D^2}{E}$
to give the χ^2 value for the particular cell of data we are considering.
- (5) Adding the χ^2 values for all the cells together to give the final χ^2 total
- (6) Determining the total degrees of freedom possessed by the components contributing to this total.
- (7) Referring to appropriate tables to find the significance "P" for a χ^2 total of this size and the appropriate degrees of freedom.

The determination of the theoretically expected frequency for the hits is simple. If no significant difference in scoring rate exists, then one would expect equal numbers of hits to be scored on each series of originals. Since the total hits = $133 + 37 = 170$, the theoretically expected frequency for each series will be $170 \div 2 = 85$. (By adopting this method, which puts the theoretically expected figure midway between the observed values, we are ensuring the most conservative estimate of the significance, and hence applying the strictest test to the telepathy hypothesis.) The calculation of the χ^2 values for the hits is then as follows:

	<u>For E.S. Originals:</u>	<u>For C.S. Originals:</u>
O	133	37
E	85	85
D	+48	-48
D^2	2,304	2,304
$\frac{D^2}{E}$	27.11	27.11
<u>Total χ^2 for the hits: $27.11 + 27.11 = 54.22$</u>		

It has been thought advisable to include the misses in the calculation as well. Here a difficulty arises because we only know the total misses: (17,440 trials, 170 hits, = 17,270 misses), but not how these were actually distributed between the two series of originals. On the null hypothesis that no difference exists, one would expect the same number of misses to be scored on each series of originals, so that the theoretically expected misses would be $17,270 \div 2 = 8,635$ misses on each series of originals.

Since we know from the first part of the calculation, that 48 more hits were obtained on Experimental Series Originals than were expected, it follows that 48 less misses were obtained than were expected. The deviation for the misses on this series is thus -48.

In the same way the deviation of misses for the

Control Series Originals will be equal in size, but opposite in sign to the hit deviation on this series, namely +48.

This method of arriving at the theoretically expected misses is somewhat arbitrary, but it has the advantage that it gives the most conservative estimate of the difference between observed and expected scores. It introduces the assumption that half of the 17,440 trials of the subjects in the Main Experiment were aimed at Control Series Originals, and half at Experimental Series Originals, whereas in actual fact the subjects presumably aimed all their 17,440 trials at the Experimental Series Originals only. The hypothesis we are testing in the calculation, however, is that chance was responsible for both hits on Experimental and Control Series Originals in the Main Experiment, and it does not seem to the writer that this hypothesis is invalidated if half the trials are allotted to each series of originals.

The calculation of χ^2 for the misses is then as follows:

	<u>E.S.O Misses</u>	<u>C.S.O. Misses:</u>
O	-	-
E	8,635	8,635
D	-48	+48
D ²	2,304	2,304
$\frac{D^2}{E}$.27	.27

$$\underline{\chi^2 \text{ total for the misses: } .27 + .27 = .54}$$

It will be seen that we have in effect used a 2 X 2 table in calculating these χ^2 totals:

E.S.O. Hits.	C.S.O. Hits
E.S.O. Misses.	C.S.O. Misses.

It remains to determine the degrees of freedom appropriate to this table. In general, the number of degrees of freedom is the number of theoretical frequencies which can be assigned arbitrarily. Our method of handling the data has been

such that a frequency entered in any one of the four cells at once determines all the others, so that this four-fold table possesses only one degree of freedom. (Dr Pratt has kindly confirmed that one degree of freedom is correct.)

It remains to refer to tables of Chi Square to find the probability value P for a Chi Square of 54.76 with 1 d.f. Reference to such a table in "Industrial Experimentation", 4th. edition, by R.A. Brownlee, (London, 1949: H.M.S.O.) page 182, gives the following:

χ^2 (1 d.f.)	3.84	P = .05
	5.41	.02
	6.64	.01
	10.83	.001

It will be seen that our value of 54.76 far exceeds 10.83, so that P is very much less than .001 (Dr. Pratt has kindly calculated that P is less than 10^{-12} .)

This means that the deviation of our observed scores from those expected on a chance basis is so great that it could be caused by chance in less than 10^{-12} experiments.

We can therefore safely reject the null hypothesis with which we started, and putting our findings more positively, say that some factor other than chance was operating to cause more hits to be scored on the Experimental Series Originals than on the Control.

Misses as well as hits have been taken into account on the advice of Dr. Greville, whom the writer consulted on the applicability of using this method of assessing the material. It will be seen however, that the misses contribute very little to the final significance of the calculation, and even if omitted from consideration entirely, would make negligible difference to the final significance. As will be seen from Table I, this holds true for all the other calculations as well. This fact is mentioned in case any reader is disposed to dispute the method of handling the misses. No matter what method is adopted in handling the misses, the

significance of the calculations will not be materially affected thereby, and the conclusions will thus still stand.

For instance, some readers might prefer to regard the Experimental and Control Series data separately, The subjects would then be assumed to make 17,440 trials aimed at the Experimental Series, and any hits scored on Control Series Originals would be regarded as misses. Next they would be assumed to have made 17,440 trials on Control Series Originals, and any hits on Experimental Series Originals would be regarded as misses. The data would then be:

Hits on E.S.O. 133, trials 17,440, misses 17,307

Hits on C.S.O. 57, trials 17,440, misses 17,403.

Using this method the χ^2 value for the misses would be half that previously obtained.

The writer did not adopt this method because he felt that it would only be justified if the judges had scored all 17,440 trials first against the 50 Experimental Series Originals, and then once more against the 50 Control Series Originals, again scoring all 17,440 drawings in this second assessment. In actual fact the judges did not do this, but scored the 17,440 subjects drawings simultaneously against all 100 originals from both series combined, so that the possibility of any one drawing scoring a hit on the Experimental Series Originals, and on the Control Series Originals was eliminated. The hits on Control and Experimental Series are in a sense related, for the more hits scored on Experimental Series Originals, the fewer drawings were left to score hits on Control Series Originals.

In any case, the significance of the hits remains unchanged, for the number of trials does not enter into their assessment. As the significance of the misses is in no case crucial to the conclusions drawn, how these misses are

handled is merely of theoretical interest.

Before giving the calculations and their results, one further point should be noted. When the theoretically expected frequency in any cell falls below 5, objectionably large errors are introduced. Consequently it has been necessary to combine B and C hits, and also G1, G2 and G3 type hits to avoid this error.

(b) Calculations C1 to C20:

Calculation
Serial No.: C 1.

Source Data
Serial No.: D 1 & 2.

Type A hits on Experimental Series Originals compared with those made on Control Series Originals to find whether the difference is big enough to be statistically significant, for combined Experimental and Control Group Subjects.

Raw Scores: all judges, uncorrected.

Observed hits on Exp. Originals:	133	Total No. of Trials:	17,440
Observed hits on Control Originals:	37	Less combined Hits:	170
Combined hits (sum)	170	Total no. of misses (difference)	17,270
Dividing by two gives expected hits for E & C Originals =	85	Dividing by two gives expected misses for E & C Originals =	8,635
Observed hits on Exp. Originals:	133	Expected misses on Exp. Originals:	8,635
Less Expected hits:	85	Subtract difference as found opposite:	+48
Difference	+48	Observed misses on Exp. Originals:	8,587
Observed hits on Control Originals:	37	Expected misses on Control Originals:	8,635
Less Expected Hits:	85	Subtract difference as found opposite:	-48
Difference	-48	Observed misses on Control Originals:	8,683

Type:	Observed No:	Expected No:	Difference "D"	D ²	$\frac{D^2}{\text{Expect. No}}$
E hits:	133	85	+48	2,304	27.11
C hits:	37	85	-48	2,304	27.11
E Misses	8,587	8,635	-48	2,304	0.27
C Misses	8,683	8,635	+48	2,304	0.27
Totals:	17,440	17,440			
Total χ^2					54.76

Degrees of freedom: 1

P (two tail) is less than .001

Calculation
checked ? Yes.

TABLE I.

SHOWING THE SIGNIFICANCE OF THE DIFFERENCES IN RAW SCORES ON THE CONTROL AND EXPERIMENTAL SERIES OF ORIGINALS FOR THE COMBINED SCORES OF EXPERIMENTAL AND CONTROL GROUP SUBJECTS FOR EACH CLASS OF HIT.

(Summary of Calculations Nos: C1 to C20)

- Notes: (1) A plus sign in front of a deviation indicates that more hits were scored on the Experimental Series Originals than on the Control Series Originals.
- (2) All Chi square totals are for one degree of freedom. Thus $\chi^2 = 3.84$ gives a P value of .05 approximately.
- | | |
|-------|------|
| 5.41 | .02 |
| 6.64 | .01 |
| 10.83 | .001 |

<u>Type of Hit:</u>	<u>No. of hits on Exp. Originals</u>	<u>No. of hits on Control Originals:</u>	<u>Difference</u>	<u>Chi square Totals:</u>
A	135	57	+ 96	54.76
B	355	342	- 7	0.074
B+C	337	350	- 13	0.26
D	114	180	- 66	15.08
E	32	126	- 94	56.42
F3	151	52	+ 99	48.84
F2	390	212	+ 178	54.50
F1	595	329	+ 266	80.86
G1	22	11	+ 11	5.66
G1, G2, and G3	28	11	+ 17	7.44
H3	80	29	+ 51	24.00
H2	220	330	- 110	22.72
H1	356	434	- 78	8.06
I	689	570	+ 119	12.12
J	153	253	- 125	40.88
K	57	39	+ 18	3.40
All hits	2,770	2,419	+ 351	35.80

Transcription checked ? Yes.

(c) Summary of conclusions drawn from Table I:

From the above table it will be seen that B, B+C, G1 and K type hits yield differences too small to be regarded as significant at the .01 level.

In categories D, E, H2, H1 and J the judges awarded more hits on the Control Series Originals than they did on the Experimental Series Originals, the differences being significant at the .01 level or better. It should be noted that these are types of hit covering fainter resemblances in shape, title or association.

For types A, F3, F2, F1, G1+G2+G3, H3, I and all types of hit combined, which include all the strict judgement categories, the subjects have scored significantly more hits on the Experimental Series Originals than on the Control Series.

The first general conclusion that emerges from a study of these figures is that some factor or factors were operating that caused the subjects to produce significantly different scoring rates on the two different series of originals. Consequently the next step will be to determine if possible the nature of the factors responsible.

(3) Factors that might account for the observed differences:(a) Pertaining to the selection of originals:

It will be remembered that one of the assumptions initially made was that the chance probabilities of drawing originals in the Control and Experimental Series were equal. It may be, however, that this assumption is wrong. It becomes necessary to review the method by which the originals were selected to decide the matter. This will involve some recapitulation.

After completing his experiment Whately Carington drew up a catalogue, giving the frequencies that different objects had been drawn by his subjects. He suggested that subsequent experimenters might use this catalogue to establish the chance probability of any object being drawn by their own subjects.

It seemed to the writer that it would be hazardous to rely on these frequencies, for many factors might operate to render them inaccurate. One of the biggest differences might be caused by variations in the subjects themselves: their ages, habits, and the culture from which they were drawn. Thus young South Africans might well draw more sporting objects than would elderly Britons. Moreover, seasonal changes would also most likely introduce changes in the frequencies. It was decided, therefore, not to use Carington's Catalogue, but rather to sample the mental content of the actual subjects taking part in the Rhodes Experiment, just before this experiment started. Thus a few weeks before the experiment was due to begin, all the subjects who had so far volunteered were asked to write down the names of 50 objects which they thought could be drawn easily and unmistakably. Helpers analysed the 15,835 suggestions returned in this "Preliminary Frequency Experiment", and forwarded to the writer a list containing titles which they believed had been suggested only once in all these 15,835 suggestions.

The writer then chose 100 titles from this list, selecting those which he thought would not be subject to seasonal changes during the course of the experiment, and which differed widely in shape, in order to minimise the possibility of confusions arising when hits came to be allocated on them. Finding titles which met these two criteria proved unexpectedly difficult, so that some of the originals finally chosen did not satisfy the writer as completely as he could have wished. (It is recommended that anyone repeating this research should give very careful consideration indeed to the originals to be used. In particular, confusions in the shape of component parts of objects are difficult to eliminate.)

Although it was intended to use titles which had each been suggested only once, errors on the part of the workers analysing the Preliminary Frequency Lists led to 7 titles being used which had each been suggested twice in the

15,835 suggestions. This was only discovered by the writer after the completion of the experiment, while carrying out a routine check of the Preliminary Frequency Material.

The 100 titles chosen were written on slips of cardboard, the cards put into a wooden box having a slit just wide enough to allow one slip to be shaken out at a time, and the box sent to the Agent with instructions to shake two slips from the box each morning, and to make target drawings of these objects. After the experiment had ended, these 50 Experimental Series Originals were sent, together with the 50 title slips still remaining in the box, to a person unconnected with the experiment. She made drawings of these 50 titles in the box, which had not been used in the experiment, these 50 drawings constituting the "Control Series Originals". She also made tracings of the Experimental Series Originals to be used by the judges. These tracings were made to hide, as far as possible, any characteristics of the agent's style, and minimise the likelihood of anyone being able to recognise a Control Series Original from an Experimental Series Original on the basis of the artists' characteristic lines, etc.

The 50 Control Series Originals, and the 50 tracings of the Experimental Series Originals were intermixed by being arranged in alphabetical order by titles, then mounted on boards, so that the judges would not be able to tell whether any particular drawing was an Experimental or Control Series Original.

The 100 drawings were then sent to the three judges who each independently compared each of the 17,440 drawings returned by the subjects with each of these 100 Originals, and allocated hits according to the instructions which they had been given.

From the above description it will be seen that the Preliminary Frequency Experiment only established what frequencies of A type hits could be expected for Control and Experimental Series Originals. We have no way of knowing

whether fainter resemblances in title, association and shape would produce equal numbers of hits on the Experimental and Control Series Originals if chance alone were operating. It is possible that the 50 originals in each series comprise a sufficiently large group to ensure that any differences in hit frequency on individual originals will even out when they are considered as a group, but we cannot be sure of this. Consequently, although the significant differences we have obtained for hit classes other than type A are more or less what we would expect on the telepathy hypothesis, they cannot be regarded as proof, or alternatively disproof of it, for we cannot be sure that the Control and Experimental Series Originals were initially equal in difficulty for these types of hit other than A.

It is suggestive, however, that deviations unfavourable to the telepathy hypothesis only tend to occur with the less strict types of assessment. When strict hit categories are examined, it is found that more hits are scored on the Experimental Series Originals than on the Control. If subsequent analyses of these types of hit other than A show secondary effects of an extra-chance nature, they can then safely be accepted as caused by the operation of E.S.P.

For the moment let us confine our attention to the type A hits. Of the 7 titles which had occurred twice in the 15,835 suggestions of the Preliminary Frequency Experiment, six found their way into the Experimental Series Originals, and one found its way into the Control Series. Hence these two series of originals cannot be regarded as being equal in initial difficulty as had been originally intended.

In the case of the Experimental Series Originals, the 50 objects were suggested 56 times in the 15,835 suggestions of the Preliminary Frequency Experiment. For the 17,440 trials of the Main Experiment the probabilities are:

$$\frac{56}{15,835} \times \frac{17,440}{1} = 61.68 \text{ hits per } 17,440 \text{ trials.}$$

In the case of the Control Series Originals, the true hit expectancy on the basis of the Preliminary Frequency Experiment will be:

$$\frac{51}{15,835} \times \frac{17,440}{1} = 56.17 \text{ hits per } 17,440 \text{ trials.}$$

This inequality in the initial difficulty of the two Series of Originals will have to be taken into account in subsequent calculations dealing with A hit frequencies.

There would be little point in recalculating the significance of the differences at this stage, as it will be necessary to make other corrections to the data to eliminate other lines of criticism. Final calculations will be presented after all corrections have been made. In the meanwhile let us turn to a consideration of the actions of the Agent, and consider whether any actions on his part could produce a spurious appearance of E.S.P. effects in the results of the subjects.

(b) Actions on the part of the Agent that might account for the observed differences.

If the selection of the Experimental Series Originals had been left to the discretion of the Agent, he might have tended to reject the titles which seemed to him to be more difficult to draw, and deliberately choose the easier ones. When the subjects came to make their drawings, they might have exercised the same type of selection, creating a false favourable deviation in favour of the Experimental Series Originals. However, in his instructions the Agent was carefully warned to accept the slips as they came from the box, and not to replace or discard any titles on the grounds that they were difficult to draw, or for any other reason. An examination of the two series of originals does not give any grounds for supposing that the selection was biased in this way. In fact the Agent specifically states that he did not deviate from the procedure laid down for him, so that it is difficult to see how any such routine and normal

actions of his during the selection process could result in the easier titles being shaken from the box and used in preference to the more difficult titles.

(ii) Unconscious selection of favourable originals due to agent's use of P.K. and pre-cognitive E.S.P.

It has been suggested by members of the Parapsychology Laboratory staff at Duke University, that the results obtained could be explained on the hypothesis that the agent was making use of Psi ability, rather than that the subjects were exercising such ability. On this hypothesis the agent would use pre-cognitive E.S.P. to become aware of which originals would tend to produce the most hits during the experiment, from chance factors, and then use psycho-kinesis to shake those slips from the selector box. Though this hypothesis was not put forward very strongly, yet it still needs to be examined, and if possible answered.

There are a number of lines of investigation that can be used to find how far this hypothesis is tenable. Firstly, we can consider whether the nature of Psi capacities as known at present makes the hypothesis plausible. It appears that in order to produce the big deviations observed, the agent's Psi ability would have to operate fairly precisely and reliably. The experimental evidence so far accumulated suggests, however, that consistently good results of this sort are the exception rather than the rule in the operation of Psi.

Secondly, let us consider what effect such a selection process on the part of the agent would be likely to have on the data. At the beginning of the experiment the task demanded of his Psi capacities would be relatively simple. There would be 100 titles in the selector box, of which let us suppose 50 are favourable, and 50 unfavourable. From chance it would be probable that one at least of the favourable titles would be lying near the slit in the selector box, so that when he came to apply his psycho-kinetic influence it would merely

have to discriminate between one or two slips at the most. As the experiment progresses, if the agent is successful in selecting favourable titles by P.K., the percentage of favourable titles still left in the selector box falls. Since there are now many unfavourable titles to each favourable, the task required of the agent's P.K. is now much more difficult, and one would expect more errors to occur. The net result would be that the greater proportion of "popular" originals (i.e. those with the greatest number of hits registered on them) would tend to be found as targets in the earlier sections of the experiment, and progressively less of them would tend to be found in the later sections of the experiment. The type A hits on the Experimental Series Originals do in fact show this tendency. The figures for the raw scores are as follows:

Throughout experiment	40 hits scored on week A Originals
	32 B
	25 C
	15 D
	21 E

Thus it would seem that the distribution of hits on the originals tends to be in line with what one would expect on this hypothesis. However, such a decline could be produced by other factors besides Psi activity on the part of the agent. For instance, if for any reason it were easier to score post-cognitive rather than pre-cognitive hits, this would produce the same decline. It must be admitted too, that the curve that the above figures produce appears suspiciously like the J-curve decline effect already so well known to parapsychologists.

All in all it appears that this line of argument cannot bring the hypothesis to a crucial test, and so will be abandoned.

For conclusive evidence that the findings cannot be accounted for on the hypothesis of the agent exercising

P.K. and precognitive E.S.P. we must anticipate findings that will be reported later in this thesis. Briefly, it will be shown that the ability of the subject to score hits seems to be related to certain personality traits possessed by **him**. This shows that the subjects were playing a vital role in the production of the hits being scored: such hits cannot therefore be attributed to actions of the agent alone. Moreover, it will be shown that the provision of links between subjects and agent increased the number of hits scored. This is of course additional evidence that some interaction between subjects and agent was responsible for the results produced.

In order to save the hypothesis, and make it explain the personality findings, one would have to assume not only that the agent could exercise efficient P.K. and precognitive E.S.P., but also that he was aware of the personality test scores of the subjects, the findings of previous investigators in this field, and the number of hits each subject would score on each original during the experiment. Even granting him omniscience in these matters, we are faced with the additional assumption that he was able to manufacture the complex relationships the data shows merely in the way he chose 50 Originals from the 100 possibilities available to him. Thus so many assumptions have to be introduced to make this hypothesis fit all the facts that it can safely be abandoned as inadequate.

(c) Factors relating to the subjects, which might influence the frequency with which certain originals came to mind during the experiment.

(i) Factors producing "Bunching", and corrections for "Bunching".

In planning this experiment it was recognised that it would be impossible to control all the factors that might influence the subjects in making their drawings. These factors might cause

the titles of the originals to come to the subjects' minds more frequently or less frequently than they had done in the Preliminary Frequency Experiment. Since Originals were chosen which had a frequency of 1, or at the most 2 in 15,835 trials, chance factors could only lessen the frequencies by that small amount per original. The greater effect of such random factors would be to increase the frequency with which certain originals were suggested to the subjects, for such factors might conceivably produce any number of excess hits on an original, as against the greatest possible negative deviation of two. In the circumstances of this experiment it would thus seem more likely that chance factors would produce a positive deviation rather than a negative one.

The fact that the subjects were almost exclusively university students in residence offered certain advantages. Thus the experimenter was able to keep a diary, in which he noted any events during the course of the experiment that might suggest any of the originals to the subjects, such as cinema pictures or posters, or other material prominently before the subjects. A visit of a ballet company to Grahamstown might cause a sudden spurt of ballet dancers to be drawn by the subjects during one particular week. It would appear likely such transient chance factors would leave their mark on the data in the form of atypical spurts or deficits of hits in certain sections of the experiment. This effect has been called the "Bunching Effect".

After the experiment had been completed, and the drawings assessed, the distributions of hits were inspected to find whether they showed any evidence of this bunching effect. With regard to the Experimental Series Originals, sixteen cases of bunching were discovered, and in many of these it was possible to relate them to some event in the environment as recorded in the diary, or to some factor in the linkage material issued to the subjects. Thus a failure in the electrical power resulted

in the subjects drawing a large number of candles and candlesticks in one particular week, but not in other weeks. However, the reactions of the subjects to any particular influence were quite unpredictable. Thus the appearance during the experiment of a poster prominently displaying Mr. Chad, one of the 100 titles in the selector box, caused the experimenter a good deal of anxiety. But although this poster remained displayed for eight days, and was repeatedly seen by a large proportion of the subjects, only one hit on Mr. Chad was made while it was being exposed.

In planning the experiment an endeavour was made to bring such chance factors under control by introducing the Control Series Originals. It was reasoned that if chance factors did play a part, they would act equally on the Control and Experimental Series Originals, and not selectively. Consequently, when hits on Control Series Originals are compared with hits on Experimental Series Originals, as has already been done, and the significance of the difference between the two groups is assessed, there would appear to be no need to correct for bunching. The fact that both the Control and Experimental Series Originals contained 50 originals each makes it likely that the chance factors operating would cancel out. However, in the case of the A hits, where we will test the significance of the deviation against the previously determined chance frequency, the bunching effect could introduce an error. Similarly, when we come to examine the data for effects produced by the introduction of linkage material, these spurts and deficits might tend to obscure basic trends in the data. It will be worthwhile then, in these two instances to make use of figures which have been corrected for bunching.

In the case of the A hits, the frequencies of hit on both Control and Experimental Series Originals have been taken into account. For other types of hit, frequencies for the Experimental Series Originals alone have been corrected, as

subsequent analyses will be concerned mainly with this data, and not hits on the Control Series Originals.

To make these corrections the following procedure was adopted. The distributions of hit for each section of the experiment were systematically inspected, and were found in general to be remarkably stable from week to week for each Original. Occasionally however, one or two of the weekly frequencies of hit on a particular original would be quite a-typical in terms of the normal score level shown in the other sections. In these cases corrections were made by averaging the scores in the "normal" sections of the experiment (but excluding the scores of week A, when no linkage material was given) and the a-typical score changed to this average level. In 15 out of the 16 corrections made to the Experimental Series Originals, the effect of these corrections was to reduce the number of hits scored on the original in question. Only in the case of "Spiffy" was the number of hits increased.

In the case of the crucial A type hits, only one suspicious bunching of hits was found. This occurred to "Candlestick" for the reasons already given. Consequently in the figures called "fully corrected A hits" only one hit on "Candlestick" has been allowed to stand, and all the rest of the A hits on this original have been eliminated for safety. (The one hit has been allowed to stand, since it would be expected on the basis of chance.)

Corrections were carried out on the hit frequencies of the following originals:

Candlestick:	A hits.
Ace of Hearts:	B and F3 hits
Rosette:	E and K hits.
Semiquaver:	B, D, F2, H3 and I2 hits.
Mr. Chad:	H1 and I hits.
Rhinoceros:	H1 hits

Spiffy:	I hits (increased).
Bishop's mitre	I hits.
Violin Bow	I hits.
Cricket Pads	I hits
Artist's Palette	I hits
Bus Stop	J hits.

Actually the frequencies as a whole were little affected by these corrections. In effect what these corrections are doing is to eliminate any hits which there is good reason to believe were produced by chance factors during the course of the experiment. The overall effect is to reduce the number of hits scored on certain of the Experimental Series Originals, but at the same time to remove any chance produced spurts that might obscure basic trends in the data, and thus clear the way for an examination of possible linkage effects.

(ii) Rational inference and Sensory Leakage
as possible hypotheses to account for the
observed results:

The subjects were kept ignorant of the use to which their Preliminary Frequency Experiment Lists were to be put, and were not allowed to see the lists compiled by other subjects. Thus they had no way of inferring from their own lists what objects might be selected for use as targets in the experiment.

Although the experimenter knew which 100 titles had been sent to the agent, he took special care not to give the subjects any indication of what these were. In any case the experimenter had no way of telling which of these titles were chosen by the agent as Experimental Series Originals, and which had been left unused in the selector box to form the Control Group Series. No leakage from him could therefore account for the results.

During the course of the experiment the agent, working 470 miles away in Cape Town, carefully refrained from writing to the experimenter. The subjects were kept ignorant of his name and address. In addition elaborate security precautions were taken by the agent to prevent anyone except himself from seeing any of the experimental material with which he was working. Thus until long after the subjects' books had been collected, nobody except the agent knew what targets had been drawn. Not even the agent knew the titles of the Control Series Originals. Consequently the hypothesis that the results can be accounted for on the basis of rational inference, or sensory leakage, or indeed fraud on the part of the subjects can thus be discarded as inadmissible, unless a flaw in the precautions taken can be found.

Here again, the fact that the data shows relationships to exist between hit-scoring ability and personality traits makes it very difficult to account for the results adequately except on the E.S.P. hypothesis.

(iii) Tendency for the subjects to score A hits on the originals they themselves had suggested in the Preliminary Frequency Experiment.

It was also suggested by Dr T.N.E. Greville that "A question might arise in the mind of the reader whether any of the hits were obtained by the same subjects who had actually suggested the objects in question when the list of suggestions was compiled. It would seem that the probabilities of a hit would be greatly increased under those circumstances and that any such hits should be ruled out."

It appears to the writer that Dr Greville's point is only valid if an analysis of hits by subjects is undertaken. If the experiment is considered as a whole, consisting of 17,440 trials, it makes no difference to the probabilities if the same subject draws an original he had suggested in the

Preliminary Frequency Experiment, since the probability of the other subjects in the group scoring hits on that original is then 0 in the total number of trials. The probability still remains 1 or 2 in 15,835 trials.

In any case, a check of the Preliminary Frequency Lists showed that in only six of the cases had the subject making the A hit originally suggested the object. Five of these hits were scored on Experimental Series Originals, one on Control Series Originals. Although the possibility of error here is thus only very small, in compiling the "Fully Corrected A Hits" these six hits have been eliminated for safety.

(d) Actions on the part of the judges which might produce the observed results:

(i) Possible bias arising from the ability of the judges to tell which were Experimental and which were Control Series Originals:

Under this heading two possibilities need to be considered.

(a) The judges might have been able to distinguish Control from Experimental Series Originals on the basis of differences in the artistic style of the two artists responsible for the two Series.

(b) The judges might have used their own Psi ability to arrive at the same distinction.

The first of these possibilities was raised by Dr. Betty Humphrey in conversation with the writer. She suggested that the two series of originals showed differences in artistic style that might make it possible for the judges to sense which originals belonged to the one series, and which to the other. Knowing this, they might consciously or unconsciously have made use of the knowledge to favour the series they believed to be the Experimental Series, by adopting a more lenient standard in awarding hits on this series than on the Control Series Originals. This would, of course, produce a spurious result.

In support of her contention Dr. Humphrey attempted to pick out from the 100 originals, drawings that she was certain had been made by the same person. As much of her own work has involved the classification of drawings into expansive and compressive types, (28, 29, 31), her attempt can be regarded as that of a highly trained expert. From the 100 drawings, she picked out 25 as having been made by the same individual, and of these 24 had actually been made by the same person. However she found no characteristics in the remaining 75 which would serve to separate them into two distinct groups.

Subsequently Mrs. E. Forster also attempted to separate the 100 originals into their two component series. She placed 26 drawings as having been drawn by the same artist in the one series, (of these 26, 17 had been drawn by the one person, and 9 by the other) and 24 drawings in the other series as having been drawn by the same individual. (Of these 24, 10 had been drawn by one artist, and 14 by the other.) She found no characteristics in the remaining 50 originals that would enable her to distinguish by whom they were drawn.

It seems safe to conclude that a deliberate study of the characteristics of the drawings might have enabled the judges to place some of the drawings as having been made by the one person, and some by another, but from the two tests reported above, it would appear that the majority of the 100 originals could not be so classified.

Dr. Pratt has also pointed out that even if it were possible for the judges to separate all the 100 originals correctly into their two series, they would still have no normal way of knowing which were Control and which were Experimental Series Originals, so that they could not be certain that they were favouring the right series of originals. The consistency of the three judges in awarding hits makes it unlikely that they were relying on guesswork of this kind.

It will be remembered that the writer himself acted as one of the three primary judges, and he would like to state that as far as he personally was concerned, he certainly made no conscious attempt to label any of the originals as possible Experimental Series Originals or possible Control Series Originals.

(b) Bias produced by the judges acquiring E.S.P. knowledge of which were Experimental Series Originals:

Another possibility that has occurred to the writer is that the judges used E.S.P. to discover which were Control Series, and which were Experimental Series Originals, and then consciously or unconsciously tended to favour the Experimental Series Originals. In all experiments involving judgements, this factor may be producing the results, and not the subjects' E.S.P. ability. This possible source of error has not been considered by many previous investigators.

This hypothesis is merely an alternative form of the previous one, so that both can be examined together.

If all the results were being produced by bias on the part of the judges, then one would expect that all the classes of hit would show a consistent deviation in favour of the Experimental Series Originals. Reference to the raw scores will show that this is not the case. In some of the less strict classes of hit, the judges have awarded more hits on Control Series Originals than on Experimental.

Moreover, the possibility of unconscious cheating is rendered less probable by the fact that three independent judges were used, and the fact that their assessments agreed with one another fairly closely. It is less likely that all three judges possessed good E.S.P. ability, and the desire to bias the results, than that one did.

The most convincing evidence against the hypothesis of bias on the part of the judges comes from a study of the A hits. If the title of the subject's drawing corresponded exactly to that of the agent's original, it was awarded an A hit. Here the judges were given very little opportunity to exercise personal judgement in what were to be classified as A hits. Minor variations in spelling might call for a decision, but in general the title was either right or wrong. As the judges worked independently, there was very little chance that all three would miss a title that could legitimately be classed as an A hit. In fact when all the subjects' drawings were classified and listed alphabetically, no possible A hits were found that had been missed by the judges. There remains the possibility that the judges had allowed some A hits that should strictly have been classified rather as B, C, D or E hits.

To check this possibility Mrs. E. Forster kindly undertook to review the A hits awarded by the judges, and to eliminate any doubtful A hits. The results of this review are given below.

(ii) Possibility that the judges adopted too lenient a standard in awarding A hits:

(a) Mrs. Forster's Review of A hits:

A total of 174 drawings were handed to Mrs. Forster for arbitration, consisting of the 170 A hits allowed by the three primary judges, plus four borderline cases rejected by them. After a study of all these possible A hits, Mrs. Forster decided that the variations they showed could best be classified under 10 headings:

Class I she described as "Letter perfect correspondence between title of drawing, and title of target." Examples are "DUMB-BELL" for "DUMB-BELL" and "PETROL PUMP" for "PETROL PUMP". This is the most rigid classification, and leaves no room for criticism whatsoever. She put 67 hits on Experimental Series Originals into this class, as against only 17 hits on

Control Series Originals.

Class II contained acceptable variants in spelling and word division, such as "WISH BONE" for "WISH-BONE" and "ACCORDION" for "ACCORDEON". She found that 21 hits on Experimental Series Originals and 6 hits on Control Series Originals fell into this class.

Class III embraced minor errors in spelling with no possibility that the title as spelled could be taken as ~~re~~ referring to any other word than the target title. Subjects' titles that had been prefixed by an "a" or an "an" were also put into this category. Examples are "BOEMERANG" for "BOOMERANG" and "A GOLLIWOG" for "GOLLYWOG". 8 hits on Experimental Series Originals and 5 hits on Control Series Originals were placed in this class.

Class IV was used for cases where the subject's title was the same as that of the target except that the subject added qualifications as to the sureness and definiteness of his impression. Examples are "POSSIBLY AN HOURGLASS" for "HOURGLASS", and "A VICE OF SORTS" for "VICE". Only two hits, both on Experimental Series Targets fell into this class.

Class V Mrs. Forster defines as "Precise target title stated after subject gives its general class. (I have followed the policy of selecting from the title one object wherever possible. These responses would not be scored hits on the more general title. Notice that no targets are compounds.)" Examples are "PLAYING CARD (ACE OF HEARTS)" for "ACE OF HEARTS", and "A MOHAMMEDAN'S HAT (I THIK ?) FEZ" for "FEZ". Here again both hits placed in this class were on the Experimental Series Originals.

Class VI has the target title given as the main response, although another object is stated following a subordinating preposition. "(Following the policy stated for class V, I would select the subordinate title as the one response.)" Only two hits, again both on Experimental Series

Originals fall into this class, namely "PALETTE WITH BRUSHES" for "(ARTIST'S) PALETTE", and "FEZ ON A MALAY" for "FEZ".

Class VII covered cases in which the target number was wrong, but only in concepts which are commonly considered in pairs, as, for example, "DUMB-BELLS" for "DUMB-BELL", and "CRUTCH" for "CRUTCHES". One hit on Control Originals, and two on Experimental Series Originals fell into this class.

Class VIII was awarded when the target title was one of two given by the subject, with a co-ordinating preposition in between. Mrs. Forster added: "All responses of this class must be considered as two. The assumption of equal likelihood of success on Control or Target title is not violated". Examples are "MOONLIT SEA AND PENGUIN" for "PENGUIN", and "ALADDIN'S LAMP (OR POT-CONTAINER)" for "ALADDIN'S LAMP". Fourteen hits on the Experimental Series Originals and two on the Control Series Originals were so classified.

Class IX comprised instances where the subject's title essentially referred to the same concept as did the target title. As examples "PENGUIN TALK" for "PENGUIN", and "NATIVE KNOBKIERRIE" for "KNOBKERRIE" can be quoted. All four hits so classified were on Experimental Series Originals.

Class X was reserved for titles which Mrs. Forster defined as "Hits which the judges have classified as A, but which I regard as definitely not belonging". Seven of the hits on the Control Series Originals, as against fourteen on the Experimental Series Originals were placed in this category. These hits rejected by Mrs. Forster were:

"FLITGUN" for "FLIT SPRAY",
 "THISTLE AND LAUREL WREATH" for "LAUREL CROWN",
 "LAUREL WREATH" for "LAUREL CROWN",
 "PATTERN" for "PALETTE",
 "CANDLE IN ELABORATE C-STICK" for "CANDLE STICK",
 "CANDLE IN STICK" for "CANDLE STICK",
 "CANDLE WITH STAND" for "CANDLE STICK",

"CANDLE HOLDER" for "CANDLE STICK",
 "BROKEN 'T' SQUARE" for "T-SQUARE",
 "WINEGLASS" for "HOURGLASS",
 "EGG TIMER" for "HOURGLASS",
 "STAGE SET TRIAL OF SAINT JOAN" for "SAINT",
 "DISCUS HURLING SEE MYRON" for "DISCUS THROWER (MODEL)",
 "JACK AND CRUCIBLE" for "CAR JACK",
 "TRIUMPHAL ARCH" for "ARC DE TRIOMPHE",
 "SPOON (violin bow)" for "VIOLIN BOW",
 "MITRE" for "BISHOP'S MITRE"
 "MITRE" for "BISHOP'S MITRE" (second instance).
 "MAN DRAWING A RICKSHAW" for "RICKSHA"
 "PRICKLY PEAR PLANT" for "PRICKLY PEAR"
 A drawing of Mr. Chad, but no title given, for "MR CHAD".

Mrs. Forster concluded: "The most rigid possible standard is letter-perfection. I would think that sub-classes I to V (128 class A hits) represent the highest practicable standard. I do not see any appreciable factor of subjectivity entering classing by these standards. All the nine sub-classes I have listed represent reasonable standards. I would analyse my hits by I to V alone, and by VI to IX alone. 21 hits placed by your judges as class A hits I would put in other classes for the reasons shown".

(b) Corrections based on Mrs. Forster's review:

Before summarising Mrs. Forster's findings, it is interesting to note that the rejected A hits in category X include a proportionate number of hits allowed by the judges on Control Series Originals, as well as those from the Experimental Series. This is additional evidence that the judges were, in fact, unbiassed, and were adopting an equal standard in awarding A hits on both series of originals.

The results of Mrs. Forster's review can be summarised as follows:

<u>Class:</u>	<u>Number of hits on E.S.O.</u>	<u>Number of hits on C.S.O.</u>
I	67	17
II	21	6
III	8	5
IV	2	0
V	2	0
Total Class I - V	100	28
VI	2	0
VII	2	1
VIII	14	2
IX	4	0
Total Class VI-IX	22	3
Class X	14	7
Total	136	38

In compiling the "Fully Corrected A hit Figures" shortly to be presented, only classes I to V above have been used, as was suggested by Mrs. Forster. This procedure again errs on the side of safety, for many of the "near misses" in classes VI, VII, VIII, IX and even X, which have been eliminated, did not appear in the Preliminary Frequency Experiment Lists, and thus could be legitimately accepted as genuine A hits.

(e) Summary of the corrections made to arrive at the "Fully Corrected A Hits":

So far we have eliminated the doubtful A hits allowed by the three primary judges, discarded the six A hits that were scored by persons who had originally suggested the titles in the Preliminary Frequency Experiment, discarded all but one of the hits scored on "CANDLESTICK" on the grounds that they

were caused by "bunching errors". The A hits which are left do not seem open to further criticism. These will be called the "Fully Corrected A Hits".

It must be stressed that the effect of all these corrections has been to discard evidence favourable to the telepathy hypothesis, not to make the hypothesis easier to establish. This will be seen from the smaller significances obtained when corrected data is used in comparison with those obtained from the uncorrected data.

Before embarking on any further calculations, however, it will be as well first to review the methods of assessment which are going to be employed.

(4) Review of further methods of assessment to be employed:

It will be remembered that so far we have merely assessed the raw score data by one method. In an attempt to account for the differences we discovered we were led to consider various factors which might account for them normally, and inevitably this led us rather far from actual assessment of the data. Before returning to such assessment, it is intended to give a brief review of the methods that will be employed, and to indicate briefly the value and disadvantages of each, before proceeding to the actual calculations themselves.

In assessing the Raw Scores, we compared the number of hits scored on the Experimental Series Originals with those scored on the Control Series, by the Chi Square method, to find whether the differences were significant. Our next step will be to apply the same method to the "Fully Corrected A Hits", in calculations C21-A and C21-B, but this time allowing for the slight initial inequality in difficulty of the two groups. By comparing the results with those of calculation C1 the effect of the corrections made can be determined, and we will be able to discover if any residue of evidence supporting the telepathy hypothesis exists after discarding any normally caused hits.

In making such an assessment we are faced with a certain difficulty in that we do not know very much about how data such as we are handling behaves: in the absence of evidence to the contrary we assume that it conforms to the normal laws of probability, and that the methods of assessment appropriate to normal material will also be legitimate to use on our data. Now the χ^2 test which we intend to employ is a distribution free test, since it does not depend on the scores being drawn from a population distributed according to the normal probability curve. But even if it yields a significant difference, this difference is only meaningful if it can be shown that the Control and Experimental Series of Originals were initially equal in difficulty. The way these two groups of originals were selected makes such an assumption plausible, but it rests ultimately on the reliability of the title frequencies established in the Preliminary Frequency Experiment. It might be argued too, that it might be naive to assume that normal probability theory applies to the kind of data we are handling (See Spencer Brown, 77). Specifically it might be suggested that the material is subject to large irregular fluctuations, and that the A hit differences observed between Control and Experimental Series Originals fall within the limits of such fluctuations, and are therefore not evidence for the existence of E.S.P. Consequently it would be as well to take as little for granted about our material as possible, and to find out as fully as we can how frequencies established in the Preliminary Experiment actually fluctuate in the Main Experiment. As a first step to this end, the hits scored on the Control and Experimental Series Originals have been compared with the theoretically expected frequencies on the basis of the Preliminary Frequency Experiment, to find whether significant variations (especially in Control Series frequency) have occurred. This is carried out in calculations C22-A and C22-B for the raw scores, and in calculations C23-A and C23-B for the fully corrected

A hits. Comparison of the results from raw and corrected scores will bring out the effect of the corrections on the data.

In the final assessment it has been thought wise to abandon all theoretical assumptions entirely, and instead determine empirically the hit totals produced in the Main Experiment on a large number of series of 50 Originals chosen in the same way as were the Control and Experimental Series. It will then be possible to assess the scores returned on the Control and Experimental Series against the behaviour of this empirically established statistical population. In this way we can determine exactly how such series of originals can be expected to behave in the Main Experiment, and hence determine whether the scores on the Control and Experimental Series depart significantly from the normal pattern.

Each of these assessments will be discussed more fully as it is presented, so we will now turn to the first of these.

(5) Calculation of the significance of the difference in number of Fully Corrected A Hits scored on Experimental as compared with Control Series Originals:

(a) Introduction:

If a significant difference is found between numbers of hits scored on Control Series Originals as against Experimental Series Originals, this can only be taken as evidence for the operation of E.S.P. if it can be shown that the difference is not due to inequalities in difficulty of the two series. As we have seen from the Preliminary Frequency Experiment, the Experimental Series included more titles which had been suggested twice than did the Control Series, so that the two series were in fact not completely equal in difficulty. Thus the problem in the calculations which follow is to allow for this initial inequality in the two series.

(b) Statistical Methods:

Two different attempts have been made to solve this difficulty of unequal initial groups:

(i) In calculation C21-A the theoretically expected hit frequencies have not been taken as equal for both Experimental and Control Series Originals, (as was done in calculations C1 to C20), but instead as proportionate to the frequencies established in the Preliminary Frequency Experiment. This still allows the observed hit frequencies in the Main Experiment to be used in determining the theoretically expected hit frequencies, This gives a more conservative estimate of the significance than if the expectancies from the Preliminary Frequency Experiment were used unaltered for the purpose.

It will be remembered that in terms of the Preliminary Frequency Experiment one would expect 17,440 trials to yield 61.68 hits on the Experimental Series Originals, and 56.17 hits on the Control Series Originals, giving a total of 117.85 type A hits.

Only 116 Fully Corrected A hits were in fact observed in the Main Experiment. If chance, and the initial unequal difficulty of the two series were the factors solely influencing the scoring of hits, one would expect $\frac{61.68}{117.85}$ of the 116 hits to be scored on the Experimental Series Originals, (namely, 60.71), as against $\frac{56.17}{117.85}$ of 116 on the Control Series Originals, (namely, 55.29). These two values have therefore been used as the theoretically expected hit frequencies.

In other respects the calculation follows the method used in calculations C1 to C20.

(ii) In calculation C21-B the Experimental Series Originals and the Control Series Originals have been equated in difficulty by reducing each series to 44 originals each of which occurred only once in the 15,835 suggestions of the Preliminary Frequency Experiment. Thus

"Vice", "Dumb-bell", "Beacon", "Penguin", "Barbed Wire Barb" and "Petrol Pump" have been discarded from the Experimental Series, because they each occurred twice in the 15,835 suggestions of the Preliminary Frequency Experiment.

By discarding "Organ" (which had also been suggested twice in the Preliminary Frequency Experiment), and five other Control Series Originals: "Accordeon", "Aladdin's Lamp", "Antbear", "Arc de Triomphe" and "Bag Pipes", (all suggested only once), the Control Series was also reduced to 44 items equal in difficulty to the Experimental Series.

Hits on these discarded originals were then also eliminated from the obtained hit totals of the Main Experiment leaving a residue of 76 Fully Corrected A hits on the Experimental Series Originals, and 23 on the Control Series Originals.

The same statistical method used in calculations C1 to C20 is now applied to this data.

(c) Calculations C21-A and C21-B:

Calculation
Serial No: C21B.

Source Data
Serial No: D4.

Comparison of the Fully Corrected A hits scored on a set of Experimental Series Originals with those scored on a set of Control Series Originals, the sets being of equal initial difficulty in terms of the Preliminary Frequency Experiment.

- (1) Only classes I to V of Mrs. Forster's A hits have been used.
- (2) Only one of the hits on "Candlestick" has been included.
- (3) Six hits made by subjects on titles suggested by them in the Preliminary Frequency Experiment have been discarded.
- (4) The following originals, each occurring twice in the suggestions of the Preliminary Frequency Experiment, have been eliminated from the Experimental series: "Vice", "Dumb-bell", "Beacon", "Penguin", "Barbed Wire Barb" and "Petrol Pump". The following originals have been eliminated from the Control Series Originals: "Organ" (occurring twice in the P.F.E.), "Accordeon", "Aladdin's Lamp", "Antbear", "Arc de Triomphe", and "Bag Pipes" (each occurring once in the P.F.E. lists.)
This leaves 44 originals of equal initial difficulty in each Series.

Data: Trials: 17,440
Hits on Experimental Series Originals: 76.
Hits on Control Series Originals: 23
Misses: 17,341.

Calculation of theoretically expected frequencies:

Total hits: 99. Hence 49.5 theoretically expected on each series of originals.

Total misses: 17,341. Hence 8,670.5 theoretically expected on each Series Originals.

χ^2 for the hits is therefore:

<u>E.S.O.</u>	Observed hits;	76	C.S.O.	23
	Expected hits:	49.5		49.5
	Deviation	+26.5.		-26.5

$$\frac{+26.5^2}{49.5} = 14.19 \text{ for each series of originals.}$$

$$\underline{\chi^2 \text{ total for hits: } 28.38}$$

χ^2 for the misses is therefore:

$$\frac{.26.5^2}{8,670.5} \text{ for each series, } = .081$$

$$\underline{\chi^2 \text{ total for misses: } .162}$$

$$\chi^2 = 28.38 \text{ (hits) } + .162 \text{ (misses) } = 28.542$$

F is less than .001

Calculation
checked ? Yes.

(d) Conclusions drawn from calculations C 21A & C 21B:

It will be seen that the χ^2 values found in these two calculations are considerably less than those obtained in calculation C 1, proving that the effect of the corrections to the data which have been carried out has been to lower the significance of the results supporting the telepathy hypothesis.

The most important fact, however, is that the calculations still show that the excess of hits scored on the Experimental Series Originals over the hits scored on Control Series Originals is highly significant even after all these deductions have been made. We concluded that some factor or factors in addition to chance are operating to produce the different scoring rates on the two sets of originals.

One would expect a significant difference of this sort if the subjects were, in fact, using E.S.P. ability to score hits on the Experimental Series Originals, which were the experimental targets, and were ignoring the Control Series Originals, which did not concern them. If it is admitted that the Preliminary Frequency Experiment was a valid method of assessing the difficulty of the various originals, and that the two series of originals were successfully equated in difficulty by its means, then it is not easy to account for the observed difference except by the telepathy hypothesis.

It might be argued by a critic, however, that the frequencies established in the Preliminary Frequency Experiment cannot be equated with the theoretical chance probabilities, as each original will naturally have its own variance. It seemed to the writer that as each series of originals contained 50 items, such differences would tend to even out, so that the previously determined frequencies for the two series considered as wholes could be taken as a fair

approximation to their theoretical chance frequency. (In this connection it should be noted that despite the highly significant difference between the hits on Experimental and Control Series Originals, the total for the Fully Corrected A hits obtained for the two series combined was 116, as against the 117.85 theoretically expected on the basis of the Preliminary Frequency Experiment. (See calculation C 21). It is odd that the hits scored on the combined 100 originals should approximate so closely to the previously determined chance frequency, whereas the frequency of hits on the two component sets of 50 originals should differ so widely from one another.

So far we have merely compared the hits scored on the Experimental Series Originals with those scored on the Control Series to find if there is a significant difference between them. It might be enlightening to compare the scores obtained on each set of originals singly with the previously determined chance frequencies as established by the Preliminary Frequency Experiment, to find the size and nature of the differences. This might help us to make a decision on the validity of the chance frequencies established by the Preliminary Frequency Experiment, and on whether the two sets of originals were successfully equated on its basis.

(6) Calculation of the significance of the difference in Raw A hits expected in terms of the Preliminary Frequency Experiment, and those actually obtained in the Main Experiment on Control & Experimental Series Originals:

(a) Introduction:

It has been thought advisable to deal first with the Raw Scores, and then with the Fully Corrected Scores, so that the effect of the corrections can be more easily seen.

(b) Statistical Method:

In the calculations following, the material for the Experimental Series Originals is assessed separately from that of the Control Series Originals. This separation makes the handling of the misses straightforward. We can assume that the full 17,440 trials were aimed first at the Experimental Series Originals, and then that the same number of trials was aimed at the Control Series Originals.

In other respects the usual χ^2 method is again used.

(c) Calculations C22-A and C22-B.

(d) Conclusions from Calculations C22-A & C22-B:

It will be seen from calculation C22-A that the subjects have tended to score more A hits on the Experimental Series Originals than would be expected on the basis of the Preliminary Frequency Experiment to a highly significant extent. This positive deviation from the previously determined chance frequency is very highly significant, P being actually less than 10^{-18} .

By contrast the subjects have scored less hits on the Control Series than would be expected on the basis of the Preliminary Frequency Experiment. This negative deviation does not succeed in reaching the .01 level of significance, but is probably significant, P lying between .02 and .01.

Interpretation of these results will be postponed until the Fully Corrected A hits have been assessed.

(7) Calculation of the significance of the difference in number of Fully Corrected A hits expected in terms of the Preliminary Frequency Experiment, and those actually obtained in the Main Experiment, on Control & Experimental Series.(a) Introduction:

By comparing the results obtained before and after making the corrections, it will be possible to determine the effect of the corrections on the validity of this type of assessment. Before embarking on the calculations, it will be as well to review the corrections which have been made to the data, and attempt to assess their probable effects.

The corrections made, have, in general, tended to reduce the number of A hits allowed on both series of originals. This will decrease the positive deviation of hits on the Experimental Series Originals from chance, while increasing the negative deviation of the hits on Control Series Originals from chance.

(b) Statistical Method:

This is the same as for calculations C22-A and C22-B, except that fully corrected A hits are used in place of Raw scores.

(c) Calculations C23-A and C23-B:

Calculation
Serial No: C23A

Source Data
Serial No: D4.

Calculation to find whether the obtained numbers of A hits made on Experimental Series Originals are significantly different from the true chance frequencies established by the Preliminary Frequency Experiment, after the following corrections have been made:-

- (1) Elimination of the doubtfully valid type A Hits. (Only classes I to V of Forster's review have been used.)
- (2) Elimination of the six hits made by subjects on titles suggested by them in the Preliminary Frequency Experiment.
- (3) Elimination of the "bunching error" on "Candle stick" by allowing only one hit to stand on that original.

The figures for the "Fully corrected A hits on Experimental Originals are:-

Expected hits on basis of Preliminary Frequency Experiment:	61.68 hits per 17,440 trials.
Obtained hits:	89.00 "
Deviation	+ 27.32 "

$$\chi^2 = \frac{(E \text{ hits} - O \text{ hits})^2}{\text{Expected hits}} + \frac{(E \text{ failures} - O \text{ failures})^2}{\text{Expected failures.}}$$

Expected failures = 17,440 - 61.68 = 17,378.32

Observed failures = 17,440 - 89.00 = 17,351.00

Deviation - 27.32

$$\begin{aligned} \chi^2 &= \frac{27.32^2}{61.68} + \frac{27.32^2}{17,378.32} \\ &= \frac{746.3824}{61.68} + \frac{746.3824}{17,378.32} \\ &= 12.1009 + 0.0429 \\ &= 12.143,8 \text{ with one degree of freedom.} \end{aligned}$$

$$\text{Hence C.R.} = \sqrt{12.143,8} = 5.485$$

F is less than .001

Calculation
checked ? Yes.

Calculation
Serial No: C23B.

Source Data
Serial No: D 4.

Calculation to find whether the obtained number of A hits made on Control Series Originals are significantly different from the true chance frequencies established by the Preliminary Frequency Experiment, after the following corrections have been made:-

- (1) Elimination of the doubtfully valid type A hits (Only classes I to V of Forster's review have been used.)
- (2) Elimination of the six hits made by subjects on titles suggested by them in the Preliminary Frequency Experiment.
- (3) Elimination of the "bunching error" on "Candle stick" by allowing only one hit to stand on that original.

The figures for the "Fully corrected A hits on the Control Series Originals" are:-

Expected hits on the basis of the Preliminary Frequency Experiment:	56.17	per 17,440 trials
Observed hits:	27.00	"
Deviation	- 29.17	"

Expected failures on the basis of the Preliminary Frequency Experiment:	17,440 - 56.17
	= 17,383.83

Observed failures =	17,440 - 27	= 17,415	per 17,440 trials.
Deviation	- 29.17		

$$\begin{aligned} \chi^2 &= \frac{(\text{E hits} - \text{O hits})^2}{\text{Expected hits.}} + \frac{(\text{E failures} - \text{O failures})^2}{\text{Expected failures.}} \\ &= \frac{29.17^2}{56.17} + \frac{29.17^2}{17,383.83} \\ &= \frac{850.888,9}{56.17} + \frac{850.888,9}{17,383.83} \\ &= 15.148,5 + 0.048,9 \\ &= 15.197,4 \text{ with one degree of freedom.} \end{aligned}$$

$$\text{Hence C.R.} = \sqrt{15.197,4} = 3.899$$

P is less than .001

Calculation
checked ? Yes.

(d) Conclusions drawn from Calculations C23-A and C23-B:

From the above calculations it will be seen that even on the most conservative scoring of A hits the subjects still score significantly more hits on the Experimental Series Originals than would be expected on the basis of the Preliminary Frequency Experiment. A deviation as big as that obtained would occur by chance in less than one out of one thousand such experiments.

Turning to the A hits on the Control Series Originals, we find that the subjects tended to produce significantly fewer hits than would be expected on a chance basis, the deviation being as significant as that produced for the Experimental Series Originals.

A comparison of these significances with those obtained for calculations C22-A and C22-B will show that the effect of the corrections has been to reduce the significance of the deviation on the Experimental Series Originals, but to increase the significance of the deviation on the Control Series Originals.

There are a number of ways of interpreting these deviations. The first is to argue that since uncontrolled factors produced a significant negative deviation in the case of the Control Series Originals, these factors should also have produced a negative deviation in the case of the Experimental Series Originals, and the fact that they did not shows that they were counteracted by the hits scored by the E.S.P. ability of the subjects. This argument welcomes the negative deviation of hits on the Control Series Originals as increasing the significance of the positive hit deviation on the Experimental Series Originals.

The second line of interpretation is in direct contradiction to the first. It could be argued that as uncontrolled factors produced a significant negative deviation of hits on the Control Series Originals,

similar uncontrolled factors could be producing the positive deviation shown by the hits scored on the Experimental Series Originals. The significant negative deviation of hits on the Control Series Originals is evidence that free drawing material is too unstable, and subject to too many extraneous influences to be assessed by the statistical method we have employed. Had the Control Series Originals shown no significant departure from the chance frequencies established in the Preliminary Frequency Experiment, then the significant positive deviation of hits on the Experimental Series Originals could have been accepted much more readily as having been brought about by the operation of E.S.P. Unless therefore we can find good reasons why such a negative deviation should be produced in the case of the Control Series Originals, we will not be able to accept the positive deviation of the Experimental Series as evidence of the subjects' E.S.P. powers. We must therefore turn to a consideration of some possible causes of such a negative deviation.

Reports of the deliberate avoidance of targets are not uncommon in the literature on E.S.P. and P.K. (56, 79). It might be that the avoidance of hits on Control Series Originals shown in this experiment can be explained in terms of psi-missing. Thus it might be argued that the subjects were aware of the 50 Control Series titles in the selector box, but since these were not used by the agent, nor linked to the subjects in any way, the subjects tended to reject such titles when they came to mind during the experiment.

Dr. T.N.E. Greville has drawn attention to a second possibility. He points out that the Chi Square procedure employed only tests the hypothesis that the results are due to chance on the assumption that the probabilities of scoring a hit were validly established in the Preliminary Frequency Experiment. He continues: "While this probability is of necessity quite arbitrary, it seems to rest on a more or

less plausible basis. Also it seems plausible to assume that this figure overestimates the probability, as it is quite likely that many of the suggestions used still would have occurred only once even if a larger number of suggestions had been obtained." This might be an explanation of the negative deviation obtained on the Control Series Originals.

Both these explanations for the negative deviation are at this stage speculative.

There is, however, a third explanation which is not speculative, and which shows that such a negative deviation on the Control Series is precisely what would be found if the subjects were successfully using E.S.P. to score hits on the Experimental Series Originals. It will be remembered that all the subjects' 17,440 guesses were ostensibly aimed at the targets in the Experimental Series. The more successful they were in scoring hits through E.S.P. on these Experimental Series Originals, the fewer guesses were left free to be determined merely by chance factors, and which could thus score hits on Control Series Originals. This will be made clearer if we consider an extreme case. Suppose that all the subjects possessed perfect E.S.P. ability. Then each guess they made would score a hit on an Experimental Series Original, and all 17,440 guesses would be hits on these originals. There would be no possibility whatsoever of scoring a hit on a Control Series Original, and hits on these would fall to zero, for the judges in making their assessments scored the subjects' drawings against all 100 Originals of both Series combined, and not in turn against each series separately. In practice of course, any type of hit, however faintly determined by E.S.P. would contribute to this effect.

Although the above explanation successfully reconciles the observed negative deviation of scores on the Control Series Originals with the telepathy hypothesis, it does not, of course, disprove the possibility that it may still be due to a random chance fluctuation. Our next assessment

will therefore attempt to discover what actual fluctuations can be expected from series of originals chosen in the same way as the Control and Experimental Series were selected.

(8) Assessment of Control and Experimental Series Originals
Scores against the actual scores obtained by a large
population of similar Series in the Main Experiment:

(a) Introduction:

So far our assessments of the data have been based on certain assumptions which the experimental design has established as plausible, namely:

(i) That the Preliminary Frequency Experiment has established sufficiently accurately the probabilities of any original being drawn from chance alone.

(ii) That the Control and Experimental Series of Originals were equated in difficulty on the basis of the Preliminary Frequency Experiment.

(iii) That chance fluctuations in hit frequency on these originals will obey the normal laws of probability, and can be assessed by the statistical techniques which have been developed for such normal material.

(iv) That χ^2 is a satisfactory technique to apply in assessing the data.

At this point it will be valuable to abandon all assumptions entirely, and investigate empirically how hits are registered in the Main Experiment on series of originals chosen in the same way as the Control and Experimental Series. We are going to make no assumptions as to the probabilities which govern such series, but instead establish what we can and can not expect from them by sampling the data actually obtained.

(b) Method:

Analysers of the Preliminary Frequency Experiment material supplied a list of 992 titles which had been suggested only once. A slip of paper was prepared for each of these 992 titles, showing how many title hits had been allowed on each

in the Main Experiment. These 992 slips were then placed in a box. (They included, of course, the 100 titles which had actually been used in the Control and Experimental Series). The writer then drew out 50 slips at random, totalled the hits scored on this series, recorded the total, returned the 50 slips to the box, shuffled the contents well, and then drew out another series of 50 slips, repeating the procedure already outlined. In this way the hits on 200 series of originals were determined.

A frequency distribution of the results was then constructed, and the mean and standard deviation determined, using the short method given in Garrett (103 .p45) The results of this investigation are given overleaf:

(c) Calculation C24:

Table II:

Calculation
Serial No: C24.

Source Data Sheets not
reproduced.

Frequency Distribution of total A Hits (uncorrected) scored
on Series comprising 50 titles each, selected at random from
all the singles titles of the Preliminary Frequency Experiment.

<u>A hit scores</u>	<u>frequency "f"</u>
75-79 111	3
70-74 111	3
65-69 1111 1	6
60-64 1111 1	6
55-59 1111 11	7
50-54 1111 1111 1111 111	18
45-49 1111 1111 1111 1111 1111 1	26
40-44 1111 1111 1111 1111 1111	25
35-39 1111 1111 1111 1111 1111 1111 111	33
30-34 1111 1111 1111 1111 1111 1111 1111 11	37
25-29 1111 1111 1111 1	16
20-24 1111 1111 1111	15
15-19 1111	4
10-14 1	1

<u>Step Midpoint:</u>	<u>f</u>	<u>'x'</u>	<u>fx'</u>	<u>fx'²</u>
77	3	7	21	147
72	3	6	18	108
67	6	5	30	150
62	6	4	24	96
57	7	3	21	63
52	18	2	36	72
47	26	1	26	26
42	25	0	+176	
37	33	-1	-33	33
32	37	-2	-74	148
27	16	-3	-48	144
22	15	-4	-60	240
17	4	-5	-20	100
12	1	-6	-6	36
	<u>200</u>		<u>-241</u>	<u>1363</u>

Assumed mean: (A.M.) 42.

i (Step interval) = 5

$$c = \frac{-65}{200} = -.325$$

$$ci = -1.625$$

$$\text{True mean} = (\text{AM} + ci) \quad 42 - 1.625 = 40.375$$

$$\text{Standard Deviation} = \sqrt{\frac{\sum fx'^2}{N} - c^2} \times i$$

$$= \sqrt{\frac{1363}{200} - .105625} \times 5$$

$$= \sqrt{6.815 - .105625} \times 5 = \sqrt{6.709375} \times 5$$

$$= 2.59 \times 5 = 12.95$$

$$\text{Standard error of S.D.} = \frac{\text{S.D.}}{\sqrt{2N}} = \frac{12.95}{\sqrt{400}} = \frac{12.95}{20} = .6475$$

Calculation checked ?
Yes.

(d) Conclusions drawn from calculation C24:

A study of the material presented in calculation C24 clears up many of the ambiguities which hindered the interpretation of previous results. It will be seen that when series of 50 originals are selected from the titles suggested only once in the Preliminary Frequency Experiment, the total hits scored on them in the Main Experiment do not show any startling departures from what would be expected in terms of the accepted probability theory. In fact the score distribution approximates quite closely to a normal probability curve.

It will be noticed that the mean score is 40.375 hits, with a standard deviation of 12.95 having a standard error of .6475. Thus it would appear that the actual hits scored on any series of originals chosen in the same way as were the Control and Experimental Series in the Main Experiment were rather less than would be expected in terms of the theoretical frequencies of the Preliminary Frequency Experiment. This may be due to any of the possible causes already discussed above, namely psi-missing, Greville's hypothesis, or the subjects' effective use of E.S.P. to score hits on the Experimental Series Originals.

The highest total of hits scored in the Main Experiment on any of the 200 series of originals sampled was 77, and the lowest total 14. It can be readily seen, therefore, that the total of 27 hits scored on the Control Series Originals is quite typical of this population, whereas the total of 89 hits scored on the Experimental Series Originals is quite atypical. Hence we are justified in concluding that the hits scored on the Control Series Originals show no deviation from what would be expected on a chance basis, whereas the surplus of hits scored on the Experimental Series Originals cannot be attributed to chance, and supports the telepathy hypothesis.

It is of course quite easy to calculate the exact probabilities for these two values. The Control Series score of 27 shows a deviation of -13.375 from the mean of the

distribution, Dividing this deviation by the S.D. of the distribution gives 1.03. Reference to appropriate tables such as in Milton Smith, (104, page 59), gives a value for P of approximately .32, which is clearly insignificant. We conclude therefore that the hits scored on the C.S.O. do not differ significantly from the empirically established chance expectancies.

By contrast the hits on the Experimental Series Originals show a positive deviation of + 48.625 from the empirically established chance frequency. This gives a value for the critical ratio $\frac{\bar{x}}{S.D.}$ of 3.75. Reference to the appropriate tables gives P lying between .00032 and .00014. We conclude that the subjects have scored significantly more hits on the Experimental Series Originals than would be expected on the basis of the empirically established mean chance expectancies.

It should be noted that the assessment we have carried out gives a conservative and not lenient test of the telepathy hypothesis, because Raw Scores were used to construct the frequency distribution, whereas Fully Corrected Scores are being assessed by its standard. In spite of the severity of this test, the results still support the telepathy hypothesis.

We are now in a position to sum up the results of this First Main Analysis. Some of these have already been presented in Table I. A summary of the results of calculations C 21A to C 24 is given in Table II overleaf:

(e) Table II:

TABLE II

SUMMARY OF THE SIGNIFICANCES OF THE DIFFERENCES
IN A HITS ON C AND E.S.O. FROM VARIOUS TYPES OF
THEORETICALLY EXPECTED FREQUENCIES.

(Calculations C 21A to C 24).

- C 21A: Significance of differences in Fully Corrected A hits scored on C. and E.S.O. using Preliminary Frequency Experiment frequencies to establish theoretically expected frequencies. P was less than .001
- C 21B: Significance of the difference in number of Fully Corrected A hits scored on two groups of E.S.O. and C.S.O. equated in initial difficulty in terms of the Preliminary Frequency Experiment Frequencies. P was less than .001
- C 22A: Significance of the positive deviation of Raw A hits from true chance frequencies established by the Preliminary Frequency Experiment. P was less than 10^{-11}
- C 22B: Significance of the negative deviation of the Raw A hits on C.S.O. from the true chance frequencies established by the Preliminary Frequency Experiment. P was less than .02
- C 23A: Significance of the positive deviation of the Fully Corrected A hits on E.S.O. from true chance frequencies as established in the Preliminary Frequency Experiment. P was less than .001
- C 23B: Significance of the negative deviation of the Fully Corrected A hits on C.S.O. from the true chance frequencies established in the Preliminary Frequency Experiment. P was less than .001
- C 24: Calculation of the Mean Raw A hits scored in the Main Experiment on a sample of 200 series of 50 Originals selected from 992 equated titles, including those used for the E.S.O. and the C.S.O.
Mean: 40.375 A hits.
S.D.: 12.95
 $\sigma_{S.D.}$.6475

Assessment of deviations of the Fully Corrected A Hits from these empirically established chance frequencies gave the following results:

E.S.O. Observed hits: 89, expected 40.375,
Deviation: +48.625 P lies between .00032 and .00014

C.S.O. Observed hits: 27 expected: 40.375
Deviation -13.375 P is approximately equal to .32

Transcription checked ?

Yes.

(9) Summary of the results of the First Main Analysis, presented in Tables I & II, and of the conclusions drawn from them.

(a) Review of the assessments & the results:

In this analysis we began by examining the Raw Score differences in hits scored on the Experimental as compared with the Control Series Originals for all the various classes of hits. These results were summarised in Table I. Reference to this table will show that in the hit categories A; F3; F2; F1; G1+G2+G3; H3 and I, and the "All Hits Combined" all showed significant positive deviations in favour of the Experimental Series Originals. This is what would be expected if the subjects were using extra-sensory perception to score hits. It will be noticed that these types of hit include all the closest resemblances in title and shape, which is again what we would expect on the basis of the telepathy hypothesis.

When hit categories covering fainter resemblances were assessed, no significant differences were found in the cases of the B; B+C; G1 and K types of hit. In the cases of D; E; H2; H1 and J types it was found that significantly less hits had been scored on the Experimental Series Originals than on the Control Series Originals. These deviations would at first sight appear to weaken the evidence in favour of the telepathy hypothesis provided by positive deviations previously found, since they would appear to have been caused by uncontrolled factors which might be responsible for producing some of the positive deviations as well. Reflection showed, however, that although they might weaken the case for considering the positive deviations in hit categories other than A as due to the

operation of extra-sensory perception, they did not weaken the evidence supplied by the A hits themselves. Only the type A hits are capable of supplying rigidly acceptable evidence for or against the telepathy hypothesis. This is so because the Preliminary Frequency Experiment only established frequencies for the A hits, and consequently we can only assume that the two groups of originals were initially equal in difficulty in terms of such title hits. For the remainder of the analysis we will therefore confine our consideration to the A hits only.

All the A hits awarded by the judges were carefully examined, and all such hits which there was reason to believe could have been caused by normal factors were eliminated. The residue of hits to which no further exception could be taken were called the "Fully Corrected A hits". Assessments were then carried out on both the Raw Scores and these Fully Corrected A hits. Three types of assessment were made of these. The results are summarised in Table II.

In the first type of assessment the hits scored on the C.S.O. and the E.S.O. were directly compared, and it was found that significantly more hits had been scored on the Experimental Series Originals than on the Control, for both Raw and Corrected scores.

In the second type of assessment the hits scored on the E and C.S.O. were compared with the chance frequencies to be expected on the basis of the Preliminary Frequency Experiment. In the case of the Raw Scores it was found that significantly more hits had been scored on the Experimental Series than could be expected on the basis of the chance frequencies, the deviation being highly significant. By contrast the hits scored on the C.S.O. were less than would be expected on the basis of the Preliminary Frequency Experiment, but this difference was only possibly significant.

When the fully corrected scores were assessed

by this method, the deviation shown by the Experimental Series remained significant, but the deviation on the Control Series was now found also to have become significant. Possible causes for this negative deviation were examined, and it was shown that it could be reconciled with the telepathy hypothesis.

At this point some doubt was thrown on the telepathy hypothesis by the fact that the Fully Corrected A Hits for both Control and Experimental Series Originals combined totalled 116, whereas the chance expectancy on the basis of the Preliminary Frequency Experiment was 117.85. It was suggested that the Preliminary Frequency Experiment might not be a completely accurate way of determining the frequencies to be expected on originals in the Main Experiment.

An investigation was consequently undertaken to find out how many hits had actually been scored in the Main Experiment on 200 series of 50 originals chosen in the same way as the Control and Experimental Series had been. It was found that the scores of these 200 series distributed normally, and when the Fully Corrected A hit Totals for the Control and Experimental Series Originals were assessed against this empirically established population, it was found that the Control Series score did not differ significantly from the chance expectancy, whereas the Experimental Series deviation still appeared as highly significant. Thus we concluded that the hits scored on the Control Series Originals did not depart significantly from chance, whereas the hits scored on the Experimental Series Originals did.

Since the mean hit expectancy was found to be 40.375 per group of 50 originals, the normal hit frequency for both Control and Experimental Series combined would be 80.75 from chance alone, under the actual conditions of the experiment, and not 117.85 as theoretically determined from the Preliminary Frequency Experiment. It would seem more valid to accept the figure obtained from actual frequencies observed in the Main Experiment, rather than the

theoretical value derived from the Preliminary Frequency Experiment.

(b) Summary of the evidence unfavourable to the Telepathy Hypothesis:

Reflection will show that this analysis has revealed three points which at first consideration seem to weaken the evidence for the telepathy hypothesis.

(i) The fact that in some of the less strict hit categories the subjects scored more hits on the Control Series Originals than they did on the Experimental Series Originals: It might be argued that as uncontrolled factors had produced these, such factors might also be responsible for some of the significant deviations in favour of the Experimental Series. Since we know that the Experimental and Control Series Originals were initially equated in difficulty for type A hits only, it might be that the differences found in the fainter resemblances were caused by the two series being of different difficulty initially for these types of hit. The observed significant deviations on the Control Series Originals suggest that we should be very cautious of drawing conclusions from hit types other than A, but any conclusions drawn from the A hits will not be affected by the ambiguity of these other results.

(ii) The fact that only 116 Fully Corrected A hits were scored on both Control and Experimental Groups combined, when 117.85 would be expected from chance alone on the basis of the Preliminary Frequency Experiment: Here a subsequent check, using actual frequencies observed in the Main Experiment showed that the actual frequency was much less, namely 80.75 hits. It seems reasonable to accept the actual frequency observed, rather than one derived from the Preliminary Frequency Experiment on theoretical grounds, which in any case commentators have shown may be inexact.

(iii) The significant negative deviation

of the Fully Corrected A hits on the Control Series Originals from the theoretically expected frequency on the basis of the Preliminary Frequency Experiment: With regard to this point it has been shown that when actual frequencies are used to assess the difference, this negative deviation becomes insignificant. It has also been shown that it is reasonable to expect some negative deviation if subjects were in fact using E.S.P. to score hits on the Experimental Series Originals.

On closer examination of each of the above points it would therefore appear that none of them are really crucial objections to the telepathy hypothesis.

(c) Summary of the evidence in favour of the telepathy hypothesis:

(i) The subjects scored significantly more A hits on the Experimental Series Originals than they did on the Control Series Originals, regardless of whether the raw or the corrected scores were used for the assessments.

(ii) The subjects scored significantly more A hits, (both raw and corrected) on the Experimental Series Originals than would be expected either on the basis of the Preliminary Frequency Experiment, or on the basis of the hit totals scored on similar series of originals in the Main Experiment.

(iii) That these positive deviations cannot be attributed to uncontrolled factors is shown by the fact that the hits scored on the Control Series Originals did not depart significantly from the actual chance frequencies observed in the Main Experiment.

(d) Conclusions:

The telepathy hypothesis can only be established if all other normal explanations for the results are eliminated. Care has been taken in this analysis to reject any hits which could possibly be caused by the operation of normal factors, and

to adopt always the most conservative tests of the significance of the data. From this treatment the results emerge as consistent with the telepathy hypothesis. However, there always remains the possibility that the writer has overlooked some factor or factors which would account for the results normally. Consequently the conclusion is stated tentatively at this stage that the data suggests that the subjects were in fact using E.S.P. to score hits on the Experimental Series Originals.

If subsequent analyses reveal other secondary effects of an extra-chance nature, then the deviations discussed above can safely be attributed to the operation of E.S.P. Let us then leave this consideration of the difference in frequency between hits on Experimental and Control Series Originals, and turn instead to an examination of the data for some other effects.

SECOND MAIN ANALYSIS:

EXAMINATION OF THE DATA TO FIND WHETHER THE
SUBJECTS TENDED TO SCORE SIGNIFICANTLY MORE
HITS ON THE EXPERIMENTAL SERIES ORIGINALS
DURING THE SECTION IN WHICH THEY WERE THE
TARGETS THAN IN OTHER SECTIONS OF THE
EXPERIMENT.

(1) <u>Introduction:</u>	159
(2) <u>Explanation of the Statistical Method:</u>	161
(3) <u>Calculations using Raw Scores:</u>	163
(a) Experimental Group Subjects: (Calculations C25 to C29):	164
(b) Control Group Subjects: (Calculations C30 to C34):	App.
(c) Table III:	165
(4) <u>Calculations using Scores Corrected for Bunching:</u>	
(a) Experimental Group Subjects: (Calculations C35 to C39):	App.
(b) Control Group Subjects: (Calculations C40 to C44):	App.
(c) Table IV	166
(5) <u>Summary of the conclusions drawn from the Second Main Analysis:</u>	167

(1) Introduction:

The method of assessment that will be used in this section was employed by Whately Carington to establish the significance of his results. In planning this investigation it will be remembered the writer tried to keep the procedure sufficiently close to that followed by Carington to justify the use of the same statistical methods.

In his investigation Carington carried out five separate experiments, I, II, III, IVa and IVb. With the exception of one subject who participated in both IVa and IVb, different subjects were used for each experiment. Carington and his wife acted as agents jointly in experiments I, III, IVa and IVb. Carington acted as agent in experiment II. In this latter experiment the procedure was also different from that followed in others, and was open to the objection that the subjects might not have obtained their results through the use of E.S.P. but might merely have been looking ahead into the content of their own minds in the future.

Carington compared the number of hits made by his subjects on the originals that were the targets in their own experiment, with the number of hits that they made on originals of other experiments at which they were not aiming. If chance alone were producing the hits they scored, then there is no reason to suppose that it would favour any particular experiment. One would expect that the hits on any particular original would be produced equally by subjects from all the experiments. He found, however, that his subjects tended to score significantly more hits on the originals of their own experiment than on the originals of other experiments. The only exception to this was in experiments IVa and IVb, which he ran concurrently. Here displacement seemed to take place between the two experiments, the subjects of IVa hitting both their own targets and those of IVb, and the subjects of IVb showing a similar effect.

In the present investigation sections A, B, C, D and E of the experiment were kept as distinct as possible both in

procedure and time, so as to correspond to Carington's five different experiments. However, in the Rhodes experiment the same subjects took part in all five sections, so that it is much more of a psychological gestalt than was Carington's experiment.

In all the statistical analyses, the weekly sections have been chosen as the units in which to work, rather than taking single days as units. This has been done for a number of reasons. Firstly, some subjects forgot to make their drawings on the proper days, and then produced a bunch of drawings at one sitting when they remembered to make them at the end of the weekly section. Since great care was taken to collect all the books for one section before the books for the next section were issued, there is no possibility that a drawing being made in one section was consciously being aimed at a target from another section. Secondly, when the writer came to examine the distribution of A hits, he found that not one of the A hits had ever been registered on the actual day on which the original in question had been one of the targets, so that there seemed little point in carrying out an analysis on such a daily basis. Consequently it was decided to examine the results to find whether they show any tendency for the most hits on Experimental Series Originals to be scored during the week in which they were the targets, than in other weeks when they were not.

This method of assessment has numerous advantages over the method employed in the First Main Analysis. If such an effect is found, it could not be due to any action on the part of the judges or the agent. There is also no need to establish that the originals are equal in difficulty since the method allows for this varying popularity. Finally, if chance factors do increase the number of hits on any original, there is no reason to suppose that the actual week in which the originals

were being used as targets would be favoured more than other weeks in which they were not the targets. For a full discussion of these points, see the comments made by Dr. J.O. Irwin, Mr. Oliver Gatty and Professor Broad, in their appendix to Carington's report. (13).

(2) Explanation of the Statistical Method used:

The technique which is going to be used can be reduced essentially to five steps. The first step is to find how many "winners" (hits on an original during the week in which it was one of the targets) were made for each type of hit. The second step is to find the number of winners to be expected on a chance basis. Thirdly the deviation of the observed winners from the expected winners is found. Fourthly we calculate the Standard Error of the expected winners, and finally obtain a Critical Ratio by dividing the deviation by the Standard Error. This C.R. is interpreted in the usual way by reference to tables.

As an illustration calculation No. C25, that for the A hit raw scores made by Experimental Group Subjects on Experimental Series Originals will be explained in detail.

For convenience the data is first arranged in a table which gives the number of hits made in each section of the experiment on the originals of section A, B, C, D and E, as shown overleaf.

By definition our "winners" are the hits made in the main diagonal: on section A originals during section A, on B originals during section B, on C originals during section C, on D during D and on E during E. Adding the figures in this main diagonal gives $2 + 1 + 4 + 1 + 2 = 10$ as the total winners observed.

The sum for each row of figures is found, and called respectively a_1 a_2 a_3 a_4 and a_5

In the same way the sum for each column of figures is found, and called respectively b_1 b_2 b_3 b_4 and b_5 .

The sum of the a's and the sum of the b's should be equal. If they do not balance, an error has been made.

Next the number of expected winners is calculated. In the example we are considering, 70 hits were made during the whole experiment. One would expect 1/5th of these to occur in each of the five weeks if they were being caused by chance. Hence to get the expected winners we divide the total number of hits made by 5. 70 divided by 5 = 14.

Next the deviation is found by finding the difference between the observed and expected winners. This gives 10 minus 14 = -4.

The next ~~step~~ step is to calculate the Standard Error, σ . This is the quantity which will serve as a measure of the extent to which the theoretically calculated expectation is likely to vary under chance conditions.

Stevens, in an article entitled "Tests of Significance for Extra Sensory Perception Data", Psychol. Rev., Vol. 46, No. 2, March 1939, has shown that the variance σ^2 is given by the formula:

$$\frac{1}{N^2(N-1)} \left[S^2(a_j b_j) + N^2 \cdot S(a_j b_j) - N \cdot \overline{S(a_j b_j)} (a_j + b_j) \right]$$

It is of course first necessary to determine whether the data is of such a nature that the formula can legitimately be applied to it. Irwin, Gatty and Broad examine this question with regard to Carington's material, and conclude that it is in order to use this method of calculating the variance. Since this investigation was planned to follow Carington's as closely as possible, it would seem safe to apply the same formula to the Rhodes data.

To calculate the variance we start by summing the corresponding row and column totals in pairs, and then multiply each of these sums $a + b$ by the previously determined product ab , so obtaining a new series of products $ab(a + b)$, which in the example we are considering has a total value of

26,940.

Next we write down the square of the sum of the ab products, and add to it the result of multiplying the sum of those products by the square of the total number of hits. From this we subtract the sum of the compound products $ab(a + b)$ multiplied by the total number of hits. Dividing this by $N^2(N - 1)$ gives the variance σ^2 , which in this example comes to 10.22

By taking the square root of the variance, we arrive at the Standard Error σ , = 3.197

Dividing σ into the deviation already found gives

$$\frac{D}{\sigma} = \frac{-4.0}{3.197} = 1.25$$

Reference to the appropriate tables (54, p 365) which we enter with $t = 1.25$, tells us that the area under the curve from $t = 0$ to $t = 1.25$ is .3944. Subtracting this from .5 gives us the probability value .1056. This is clearly not significant, as chance could produce a deviation of this magnitude in roughly 21 out of 100 such experiments. (The probability value .1056 is doubled, since the direction of the deviation we are testing is not considered.)

Since the results of the Experimental Group Subjects might have been more favourably affected by the linkage factors introduced than the results of the Control Group Subjects, it has been thought advisable to examine the results of each of these groups of subjects separately. There is, of course, no point in considering the hits made on the Control Series Originals: the calculations will deal solely with the hits made on Experimental Series Originals.

The assessments have been carried out first on the raw scores, and then on the scores that have been corrected for bunching. It was thought that trends in the raw scores might be obscured by these chance bunching errors.

Calculations for the various categories of hit follow:

- (3) Calculations using raw scores: C25 to C34:
- (4) Calculations using scores corrected for bunching: C35 to C44.

Calculation
Serial No.: C 25.

Source data
Serial No.: D 5.

Calculation to find whether the "winners" differed significantly from chance expectation.

Type A hits on Experimental Series Originals made by Experimental Group Subjects. (Raw scores, all judges).

		Sections of Exp. in which hits were made:					Totals:
		A	B	C	D	E	
On Originals of sections:	A	2	3	4	5	9	23 = a_1
	B	6	1	3	1	4	15 = a_2
	C	0	3	4	3	3	13 = a_3
	D	0	1	3	1	3	8 = a_4
	E	1	2	4	2	2	11 = a_5
Totals:		9 = b_1	10 = b_2	18 = b_3	12 = b_4	21 = b_5	70 = N

Diagonal total
(Observed Winners) = 10 = O_W

$a_1 b_1$	207	$a_1 + b_1$	32	$a_1 b_1 (a_1 + b_1)$	accumulating:
$a_2 b_2$	150	$a_2 + b_2$	25	$a_2 b_2 (a_2 + b_2)$	
$a_3 b_3$	234	$a_3 + b_3$	31	$a_3 b_3 (a_3 + b_3)$	
$a_4 b_4$	96	$a_4 + b_4$	20	$a_4 b_4 (a_4 + b_4)$	
$a_5 b_5$	231	$a_5 + b_5$	32	$a_5 b_5 (a_5 + b_5)$	
S_{ab}	<u>918</u>	S_{a+b}	<u>140</u>	$S_{ab}(a + b)$	<u>26,940</u>

$$E_W = \frac{N}{5} = \frac{70}{5} = 14 \quad O_W - E_W = -4 = D$$

$$N = 70 \quad N^2 = 4,900 \quad N^2(N - 1) = 338,100$$

Whence: $S_{ab}^2 = 842,724$

$$N^2 \cdot S_{ab} = 4,498,200$$

$$\text{Sum: } 5,340,924$$

$$N \cdot S_{(ab)(a+b)} = 1,885,800$$

$$\text{Subtracting: } 3,455,124$$

Dividing by $N^2(N - 1)$ gives: $10.22 = \sigma^2$

Whence $\sigma = 3.197$

So that $\frac{D}{\sigma} = \frac{-4}{3.197} = -1.25$

P two tail is less than:
(see table).

Calculation
checked? Yes.

TABLE III

SUMMARY OF MATERIAL BEARING ON WHETHER SUBJECTS TENDED TO SCORE MORE HITS ON ORIGINALS DURING THE WEEK IN WHICH THEY WERE THE TARGETS, THAN IN OTHER WEEKS.

RAW SCORES ON EXPERIMENTAL ORIGINALS ONLY.

(A) EXPERIMENTAL GROUP SUBJECTS:
(Calculations C25 to C29 only)

<u>Type of hit:</u>	<u>Observed</u> <u>hits</u>	<u>Expected</u> <u>hits</u>	<u>Deviation</u>	<u>D</u> <u>O</u>
A	10	14	-4	1.25
A + B + C + D + E	49	62.8	-13.8	2.04
F2, F3, + G2, G3, + H2, H3	82	82.6	-0.6	0.07
All F, G and H.	176	177	-1.0	0.08
I + J + K	105	93	+12	1.43

(B) CONTROL GROUP SUBJECTS:
(Calculations C30 to C34.)

A	10	12.6	-2.6	0.83
A + B + C + D + E	67	60.4	+6.6	0.96
F2, F3 + G2, G3 + H2, H3.	95	86.8	+8.2	1.00
All F, G and H	196	187	+9.0	0.75
I + J + K	97	82.8	+14.2	1.81

None of the above reach the .01 level of significance.

Transcription
checked ? Yes.

TABLE IV:

SUMMARY OF THE MATERIAL BEARING ON WHETHER SUBJECTS TENDED TO SCORE MORE HITS ON ORIGINALS DURING THE WEEK IN WHICH THEY WERE THE TARGETS, THAN IN OTHER WEEKS.

SCORES CORRECTED FOR BUNCHING.
HITS ON EXPERIMENTAL SERIES ORIGINALS ONLY

(A) EXPERIMENTAL GROUP SUBJECTS
(Calculations C35 to C39).

<u>Type of hit:</u>	<u>Observed hits</u>	<u>Expected hits:</u>	<u>Deviation</u>	<u>D</u> <u>0</u>
A	10	13.2	-3.2	1.06
A + B + C + D + E	45	53.6	-8.6	1.36
F2, F3, + G2, G3 + H2, H3	76	79.4	-3.4	0.43
All F, G + and H	170	169.4	+0.6	0.05
I + J + K	80	75.2	+6.8	0.91

(B) CONTROL GROUP SUBJECTS:
(Calculations C40 to C44).

A	10	11.8	-1.8	0.59
A + B + C + D + E	55	54.8	+0.2	0.03
F2, F3 + G2, G3 + H2, H3.	87	84.2	+2.8	0.25
All F, G and H	188	181	+7.0	0.60
I + J + K	90	79.6	+10.4	1.37

None of the above reach the .01 level of significance.

Transcription
checked? Yes.

(5) Summary of conclusions drawn from the Second Main Analysis

Reference to tables III and IV will show that all of the differences tested fail to reach the .01 level of significance. This analysis has therefore produced no additional evidence that would strengthen the telepathy hypothesis. However, should this hypothesis be established by later analyses, this Second Main Analysis will tell us a good deal about the way in which the subjects exercised their E.S.P. ability. It shows that the subjects' hits scored on any particular original were scattered throughout the whole experimental period, there being no significant increase in the number of hits scored on such an original during the week in which it was actually being used as a target.

Dr. Rhine has pointed out that this finding has very important implications for future research. In assessing their results most investigators have hitherto looked for hits on the direct targets at which the subjects were aiming, or at most, displacement hits to a nearby target. In the Rhodes experiment the hits appear to be dispersed equally throughout the whole experimental period, and not centered in particular weeks.

It is interesting to note that Carington also found much the same effect when he carried out experiments IVa and IVb concurrently, though using different subjects. His subjects tended to score hits on the originals of both experiments, rather than only on the targets of their own experiment. This was not the case with the other experiments in his series.

In explanation of the Rhodes results, it is suggested that although an attempt was made to keep the sections of the experiment as distinct from one another as possible, yet the subjects thought of these sections as parts of a whole programme which they had set themselves to complete. Thus the psychological gestalt produced tended to make for hits displaced over the whole period. Had it been possible to use different subjects for each section, the type of effect found by Carington might have been reproduced in the Rhodes data.

THIRD MAIN ANALYSIS:PERSONALITY.

(1) <u>Introduction.</u>	171
(2) <u>Review of the research findings on the relationship of personality differences to E.S.P. ability:</u>	172
(a) Expansion-Compression ratings.	173
(b) Subjects' attitudes and interests.	176
(c) Work on Interest Inventory ratings.	177
(d) Introversion-Extraversio ratings.	179
(e) Sheep-Goat separations.	179
(f) Ratings of psychological adjustment.	181
(g) The relation of E.S.P. ability to psychotic and neurotic syndromes.	184
(h) Conclusions.	186
(3) <u>Predictions as to what trends should be shown by the Rhodes data if it was produced by E.S.P. ability on the part of the subjects.</u>	187
(4) <u>The tests used in the Rhodes investigation: their validity and reliability:</u>	188
(a) The Bernreuter Personality Inventory.	188
(b) The Pressey X-0 Tests.	189
(c) The Allport-Vernon Study of Values.	190
(d) The Washburne Social Adjustment Inventory.	191
(e) The Psychic Questionnaire.	193
(5) <u>Statistical Methods: the selection of suitable techniques.</u>	194
(6) <u>Product-moment correlations of subjects' total hit scores (all types) with their personality test scores:</u>	
(a) Introduction.	199
(b) Statistical method.	199
(c) Calculations C45 to C58.	201
Table V: Correlations obtained between subjects hit scores & their personality test scores.	203
(d) Conclusions drawn from Table V.	204

THIRD MAIN ANALYSIS:PERSONALITY.

(7) <u>Examination of the mean personality score differences of groups of subjects classified in terms of A hits scored:</u>	205
(a) Introduction.	205
(b) Statistical method.	206
(c) Calculations C55 to C68.	208
Table VI: Personality test score means for various subject groups.	210
(d) Conclusions drawn from Table VI.	211
(e) Table VII: Significance of the mean score differences shown in Table VI.	214
(e) Conclusions drawn from Table VII.	215
(f) Calculations C69 to C72.	215
Table VIII: Bernreuter B3-I mean score differences in more detail.	216
(g) Conclusions drawn from Table VIII.	217
(h) Calculations C73 to C76.	217
Table IX: B3I Introversion scores of groups scoring A hits compared with those of groups suggesting titles in the Preliminary Frequency Experiment.	218
(i) Conclusions drawn from Table IX.	219
(8) <u>Nonparametric assessment of the personality test scores giving significant differences by the mean difference parametric method: Bernreuter B3-I and F1-C scores:</u>	220
(a) Introduction.	220
(b) Statistical method.	220
(c) Calculations C77 to C80.	223
Table X: Median test results of F1-C and B3-I score differences for various subject groups.	225
(d) Conclusions from Table X.	226
(9) <u>Brief summary of the findings so far obtained by the three methods of assessment used: product-moment, mean difference, and median test methods.</u>	228

THIRD MAIN ANALYSIS:

170.

PERSONALITY:

(10) <u>Sheep-Goat ratings from the Psychic Questionnaire:</u>	229
(a) Introduction.	229
(b) Statistical methods.	230
(c) <u>Graph No: 1</u> Scoring rates for Sheep, Goat & Openminded subjects	233
Calculations C81 to C83.	232
(d) Conclusions from the Sheep-Goat separations.	239
(11) <u>Combination of Sheep-Goat ratings with adjustment ratings:</u>	240
(a) Introduction.	240
(b) Statistical method.	240
(c) Calculations C84 and C85.	241
Graph No: 2 Scoring rates for well, poor & averagely adjusted subjects	244
Calculation C86.	245
(d) Comments on the preliminary adjustment separations.	247
(e) <u>Graph No: 3</u> Scoring rates for 9 categories of combined Sheep-Goat & adjustment subject groups.	248
Calculation C87.	249
(f) Comments on the combined Sheep-Goat and adjustment separations.	252
(12) <u>General conclusions from the Personality Test Material in terms of the predictions made.</u>	253
(13) <u>Comparison of the answers to selected questions from the Psychic Questionnaire returned by a subject group scoring 2 or more A hits, with the answers returned by a subject group failing to score A hits, to find whether the differences revealed are statistically significant.</u>	256
(a) Introduction.	256
(b) Statistical Method.	256
(c) Calculations C88 to C103.	259
Table XI: Answer frequencies for groups containing all subjects.	261
Table XII: Answer frequencies for groups of subjects returning 50 drawings.	263
Table XIII: P values of tables XI & XII compared.	265
(d) Conclusions from Tables XI, XII and XIII.	266

(1) Introduction:

It has been pointed out by Gardner Murphy, (45) that in looking for personality correlates with E.S.P. ability we are reasoning from the fact that normal sensory perception shows such relationships, and are assuming the E.S.P. is similar to S.P. in this respect. Kahn, (59) points out that there is no empirical reason to believe that E.S.P. expresses the dynamics of personality, although this has been postulated by Bell (5) and others on theoretical grounds.

Despite the possibility that the search for personality correlates may prove fruitless, there is still considerable value in looking for such relationships, since if they are found, they will increase our knowledge and control of the E.S.P. process, and will enable us to make predictions as to the type of E.S.P. performance to be expected from particular personality types. Once such relationships have been demonstrated empirically, they can be used also to test any new data produced. If such data is genuinely the result of E.S.P., then it should show the same relationships as have been established by previous investigators. Here then is an additional test that can be applied to the Rhodes material which may help to establish whether or not it has been produced by E.S.P. ability on the part of the subjects. It is proposed to survey the research findings bearing on personality in relation to E.S.P.; predict what relationships can be expected in the Rhodes data if it is due to E.S.P., and then proceed to examine the data to find whether it bears out the predictions made.

A second justification for looking for personality relationships in the Rhodes material is that such relationships would be quite relevant to the linkage problem, which is the main concern of the present research.

(2) Review of the research findings on the relationship of personality differences to E.S.P. ability.

Many of the earlier E.S.P. researches had suggested that good and bad scoring was in part produced by the personality factors in the experimenters and subjects. It was thus to be expected that sooner or later workers would follow up these incidental observations by more systematic study of such personality differences, especially as many techniques and tests for measuring personality had already been developed by psychologists.

Among the psychological tests that have been, or are being used by researchers in relation to E.S.P. are the following:- The Stuart Interest Inventory; the Kulder Personal Vocational Preference Record; R.B. Cattell's 16 Factor Inventory; the Guilford Martin Inventory of factors S T D R C G A M I N; the Minnesota Multiphasic Inventory, The Maslow Security-insecurity test; the Bernreuter Personality Inventory, the Heston Personal Adjustment Inventory; projective tests like the Rorschach, the Rosenzweig Picture Frustration Test; the assessment of the form of subjects' drawings, their handwriting, and relaxation, as well as intelligence tests of various kinds.

Hand in hand with the investigation of the normal personality has gone study of severely disrupted states. Thus accident concussed patients, brain damage patients, and psychotics have also been studied to find how E.S.P. is affected by these syndromes.

Although many investigators have contributed to the advancement of knowledge in these fields, the most systematic brilliant and sustained contributions have come from three workers: Dr Stuart, (80 to 88) Dr. Humphrey (27 to 37, 76, 88), and Dr. Schmeidler (60 to 70). The review that follows deals largely with the techniques they pioneered, and the findings which these techniques have yielded.

Though the findings themselves are relatively clearcut, a possibility of confusion arises because of the variety of experimental techniques used. Thus a relationship that is reported for clairvoyant conditions may be reversed under G.E.S.P. conditions. This makes it necessary for the procedure used by the experimenters to be constantly born in mind.

Secondly, many of the traits sampled by psychological questionnaires cannot be considered basic units of personality, but may themselves be made up of a complex of factors. Although an investigator may report a perfectly valid relationship between such a complex trait, and E.S.P. scoring, there is no way of knowing what factors in what weightings are actually determining the personality scores of the subjects. Consequently a different investigator, working with the same trait, may fail to find any confirmation of the relationship, since in his subjects different weightings of the component factors may be contributing to produce the personality score. This lack of precision in the basic measuring devices renders the interpretation of even well-established findings difficult.

Finally, it will be found that many of the relationships reported relate to the response patterns of the subjects and the way they use their E.S.P. ability, and not to the amount of E.S.P. ability shown.

If these points are borne in mind, we can proceed to outline some of the better established findings.

(a) Expansion-Compression Ratings:

Expansion-Compression was a concept first developed by Elkisch (19) as a means of classifying childrens' drawings to reveal neurotic trends. She defines her criteria for this classification as follows:-

"Expansion is expressed (a) through a widening (opening) of the space at the drawer's disposal, by presenting only part

of the object --- this might also be an abstract design --- which has to be completed by imagination; (b) through the creation of a "spacious" background which may be presented on a piece of paper of any size; (c) by the creation of an experience of space by means of rhythm and integration. For example, a well formed representation of an explosion expands the space by "bursting" it. Such an expression conveys controlled aggressiveness, wilful and forceful activity; (this is rather a virile form of expression). Expansion stimulates the imagination dynamically. It conveys an atmosphere of freedom, courage, adventure, and may be a symptom of vitality and healthily developed extraversion."

"Expansion stands for a direction towards the surrounding world; for the potential ability of making contact."

"Compression is based on a meticulous, fearful concept of space, expressed either in the spatial appearance of the object itself, or in its spatial relationship to other objects, or to the space at the drawer's disposal. Compression conveys a feeling of discomfort, of being shut in, of pressure and compulsion. Compression may be, if connected with other traits, a symptom of neurotically developed introversion, even of a compulsion neurosis. Compression stands for isolation."

Besides Expansion-Compression Elkinsch described three other qualities: Rhythm-Rule, Complexity-Simplicity and Integration-Disintegration.

The Expansion-Compression classification was applied by Humphrey to drawings produced by subjects taking part in E.S.P. experiments in order to find whether their E.S.P. performance was related to the form of the drawings produced. (28, 29, 31).

She found that a separation of the high-scoring from the low-scoring subjects was possible on the basis of expansion-compression ratings. In clairvoyant experiments she reported that subjects rated as expansive gave scores above chance,

while the compressives scored below chance expectation. Examination of the way in which the compressives were scoring showed that their low scores were not caused by lack of E.S.P. ability, but that this ability was producing significant backward displacement effects, instead of hits on the actual targets.

In G.E.S.P. experiments the Expansion-Compression criterion was still successful in separating high-scoring and low-scoring groups, but this time the compressives scored positively, and the expansives negatively. (29).

Subsequently Bevan (8) reported on the results of tests carried out in light and darkness, which tended to confirm Humphrey's findings. Here the compressives did well in half the tests, when the sender was looking at the pictures, and badly in the other half, when no-one was looking at the targets.

A review of the subsequent research on the Expansion-Compression rating was published by Humphrey in July 1951. (36). This reported that a total of twelve clairvoyant series had been evaluated by this technique. In ten out of the twelve series the expansives obtained a higher score than the compressives.

The evidence supported the view that the Expansive-Compressive quality expressed the subject's mood of the moment, rather than a stable personality factor.

In the same period nine G.E.S.P. series had been evaluated, and in these the compressives tended to make higher scores than the expansives in eight out of the nine cases.

As far as card-calling experiments were concerned, she reported that "In general then, we may sum up the card-test results by saying that there is some tendency for expansive subjects to score positively on clairvoyant card tests, and for compressives to score below chance expectation. In the 29 experiments 17 followed the trend just stated, 9 showed a reversal, and 3 showed no difference at all

between the two classifications of subjects. Thus the expansion-compression measure is by no means as effective in separating high and low scorers in card tests as it was in ESP drawing tests."

Considering clairvoyant card tests, she concludes: "the expansion-compression rating is not dividing subjects according to whether they will score positively or negatively as we might suppose from looking at the earlier graph. Rather it appears to be separating subjects as to the type of hit distribution they will give."

A recent research undertaken by West (99) with slight modifications in experimental procedure, failed to find any evidence to support Humphrey.

A recent research, reported by Kahn (39) found a difference between expansives and compressives in the direction predicted by Humphrey, but this was too small to be significant.

(b) Subjects' attitudes and interests:

This corner of the personality field has also proved popular with investigators. In 1945 Grela (23) reported an investigation into the effect of hypnotically induced attitudes on the scoring of E.S.P. subjects, that suggested that the highest scoring was produced during sessions when the subjects were inspired to believe in E.S.P. and their ability to score well, and the lowest scoring in sessions when they received negative suggestions intended to instil doubt.

MacFarlane (42) found that when two agents were used simultaneously, his subjects tended to score hits on the pack of cards being used by the agent who had previously produced good results, rather than on the pack being used by the agent who had previously been unsuccessful.

Soal and Pratt (79) found that the attitude of the subject towards the test tended to influence the manner in which hits were scored.

There is thus considerable support for the view that the motivation of the subjects influences the way in which they use their E.S.P.

(c) Work on Interest Inventory Ratings:

In 1946 Stuart (87) reported work with an Interest Inventory consisting of sixty items such as: Tennis; History; Pin-ups; Children, etc. Next to each item was a five-point scale, ranging from "Like very much" to "Dislike very much". The subject used this scale to indicate his degree of like or dislike for each item. Stuart used drawings in sealed envelopes as target material.

Stuart found no evidence to suggest that the subjects tended to avoid targets that they disliked, or score their hits on targets that they liked. However his inventory provided a general measure of the affectability of his subjects --- that is, the extent to which they generally liked or disliked things. To get this score Stuart converted his subjects' likes and dislikes into numbers by using the formula LL = 2; L = 1; I = 0; D = -1; and DD = -2. When the test was scored in this way it yielded a distribution ranging from -12 to +78 with a median of +55.5. He then classified his subjects into a "Midrange" and "Extreme" group, which Humphrey describes as follows:-

"Midrange subjects were those whose inventory scores did not differ appreciably from the average score of their group; extreme subjects on the other hand, were those whose scores were either high or low in relation to the average group score."

Stuart found that these two groups showed differences in E.S.P. performance. The "Midrange" group scored positively on the actual targets at which they were aiming, and his "Extreme" group negatively. Breaking down his data revealed that his "Extreme" group were scoring negatively because they were displacing their hits onto adjacent targets. This suggests

that the more affectable the subjects, the less control they had over aiming their E.S.P.

Stuart also derived a second measure of affectability from his inventory scores in the following way: "Since affectability may be thought of as related to the number of likes and dislikes expressed by the subjects, we may count all likes and dislikes as of equal weight. Or more simply, we may count the number of items marked "I" for "indifferent". The affectable person would be indifferent to few items, the unaffected, to relatively many." When the subjects were separated into two groups on this basis, their scores showed the same trend as revealed by the other measure of affectability.

After Stuart's death in 1947, Humphrey continued his work, and reported the results of an analysis of eight E.S.P. series which Stuart had carried out, and an additional eight series which she had analysed on the same lines. The pooled results supported Stuart's findings. (32)

Humphrey proceeded to extend the investigation by combining Stuart's "Midrange-Extreme" ratings with her own "Expansion-Compression" ratings. She found that expansive-midrange subjects scored positively, and compressive- extremes negatively. The difference between the average scores of the two groups resulting from the combination of the two personality measures was greater than that obtained when either measure was used separately. (30)

This work was followed in 1950 by the report of a new scale for separating high-scoring and low-scoring subjects in E.S.P. tests. (34). This new scale was derived from the Stuart Inventory, and consisted of the 14 items which high-scoring and low-scoring subjects answered differently most consistently. This scale was shown to be more effective than the original full-length version.

In the same year Humphrey (35) combined midrange-extreme, expansive-compressive and the 14-item interest ratings to find whether this combination would be more effective in separating high-scoring and low-scoring subjects. She found that subjects rated "expansive-midrange-high" scored positively, above chance expectation, while the "compressive-extreme-low" group scored below chance.

By July 1951 Humphrey had analysed 32 experimental series in all, embracing nearly 900 subjects. She found the results confirmed previous findings: midrange subjects scored more hits than would be expected from chance, whereas the extreme subjects scored fewer than the expected number of hits. The difference between the two groups was significant at the .0004 level. (36)

(d) Introversion-Extraversion Ratings:

Introversion-Extraversion ratings given by the Bernreuter Personality Inventory (7) have been shown by Humphrey (37) to be effective in predicting high-scoring and low-scoring subjects in E.S.P. tests. Using the data from three G.E.S.P. experiments for her analysis she found that extraverts scored significantly above chance, whereas introverts scored significantly below, the difference between the average scores of the two groups being significant at the .005 level.

(e) Sheep-Goat separations:

A study by Pratt and Price (49) suggested that attitude and mental set were important in determining the E.S.P. scoring level. Consequently Schmeidler proposed the hypothesis that "when tested by equally rigid procedures those believing in the possibility of E.S.P. should give themselves to the task with greater freedom, and so make higher scores, than those rejecting its theoretical possibility". This hypothesis she proceeded to test in a series of researches (60 to 70).

Before beginning any experiment, Schmeidler asked her subjects to indicate whether they classed themselves as "sheep" (those who accepted the possibility of E.S.P. occurring) or "goats" (those rejecting the possibility that E.S.P. could occur.) She then proceeded to examine the scores of the two groups to find whether they showed any significant differences.

During the period 1943 to 1945 she reported seven experiments (60, 61, 62 and 70) the data from which showed that higher average E.S.P. scores were obtained by sheep than by goats, though the total difference in average score was quite small.

Since 1945 Schmeidler has confirmed her findings with a further series of experiments. Here the differences in means between the two groups were smaller, but still in the predicted direction. (70)

In an experiment with 10 sheep, 10 goats and 10 undecided subjects Bevan (9) found that his undecided group obtained the highest score, next highest were the sheep, and lowest the goats. The believers were more successful at clairvoyance, but the undecided subjects and the disbelievers did better at G.E.S.P. However the small number of subjects used make Bevan's findings of little real value.

Kahn (39) has also latterly attempted to test the sheep-goat findings, but has obtained no conclusive results, and has concluded that Schmeidler has not stated her sheep-goat hypothesis in a form that can be clearly repeated. This conclusion seems to the writer surprising, since the hypothesis would appear to be quite simple and unambiguous as set out.

A further attempt to test sheep-goat predictions has been made by Casper (15). The only significant difference

he found was between the sheep, who scored positively, and the indecisives, who scored negatively.

In reviewing the findings from sheep-goat investigations in a report published September 1950, Schmeidler points out that although the effect of the separation is small, it can be increased by further classifying the sheep and goats as well-adjusted and poorly-adjusted. (69) Consequently further discussion will be postponed until the work on adjustment ratings has been considered.

(f) Ratings of Psychological Adjustment:

Although we will be concerned mainly with the work of Schmeidler in this section, it is first necessary to outline some of the findings of other workers as a background.

Smith and Humphrey (76) found that high E.S.P. scorers were rated "secure" on the Maslow Security-Insecurity test, whereas those rated "insecure" tended to score low.

An exploratory correlation study by Humphrey (27) using the Bernreuter Personality Inventory produced small plus and minus correlations in the neighbourhood of .2 or .3 when the subjects' E.S.P. scores were correlated with their personality scores on the various scales.

Rivers (58) working with the Mental Health Analysis (Thorpe, Clark and Tiegs) found no significant results.

Schmeidler, in a series of experiments, (reported in 65, 67, 69), investigated the relationship of Rorschach ratings to E.S.P. scoring levels. She found that

- Well-adjusted sheep scored well above chance.
- Poorly adjusted sheep scored just above chance.
- Poorly adjusted goats scored just below chance.
- Well-adjusted goats scored well below chance.

The Rorschach records from two batches of subjects, 250 and 529 in number, were analysed, and it was found that

seven Rorschach "signs" appeared much more frequently in the records of subjects making poor E.S.P. scores than in the records of subjects making good scores. ("Good" scores were taken to mean scores above chance for the sheep, or scores below chance for the goats.) These signs Schmeidler lists as follows:-

R+	(More than 30 responses in a group test, More than 50 responses in an individual test.)
F%+	(50% or more of the responses determined primarily by form.)
IR	(More than half of the human movement responses characterised by rigidity.)
Total Movement ++	(60% or more of all responses showing movement or expression.)
No shock	(Neither colour nor shading shock even when both are scored very lightly.)
CF+	(2 or 3 strong or whole colour-form responses if there is no more than one main form-colour response; or at least twice as many colour-form responses as form-colour responses.)
C+	(2 or more colour description or colour symbolism responses; or one or more pure colour responses.)

Schmeidler concluded:

(1) There was a tendency for the attitudes of the subjects to affect their scoring level.

(2) This tendency could be obscured by the operation of other factors.

(3) Subjects who were well-adjusted gave less equivocal data than subjects whose adjustment was poor.

(4) A cold withdrawn attitude mitigated against good E.S.P. scores, as did marked impulsiveness and lack of

emotional control, or excessive, almost compulsive, mental activity.

The same investigator has also used the Rosenzweig Picture Frustration Test with 446 E.S.P. subjects, taking a clairvoyance type experiment. (68). She found a significant positive relationship between E.S.P. success and "Impunitiveness" and a significant negative relationship between E.S.P. success and "Extrapunitiveness", thus confirming a finding that had previously been reported by Eilbert (18). This finding she interprets in general to mean that "in general, subjects who are ready to be tolerant and friendly will, under these conditions, have more E.S.P. successes than subjects who are ready to show overt hostility"

The work of Bates and Newton (4) with mental patients tends to confirm this finding. They classified patients as co-operative, apathetic, or irritable, and found that their co-operative patients were the most successful, scoring highest in high aim experiments, and lowest in low aim experiments. The apathetic subjects were next, and the irritable subjects were the least successful.

In a recent investigation Kahn (39) tested the hypothesis that "subjects who are rated as well adjusted will tend to score above chance, whereas subjects who are rated as poorly adjusted will tend to score below chance." As a measure of adjustment he used the Heston Personal Adjustment Inventory. (24). It should be noted that this hypothesis is more in line with Humphrey's findings than Schmeidler's, since it takes no account of the belief of the subjects in E.S.P. While Kahn's results were in the direction predicted by Humphrey, they did not reach the .01 level of significance.

In conclusion, the evidence on adjustment so far accumulated from psychometric studies suggests that it is not directly connected with the amount of E.S.P. ability which a subject possesses. It would seem to be connected

rather with the way they use that ability. Well-adjusted subjects appear to be able to score consistently above or below chance, the direction of their scoring being determined by their attitude to E.S.P., whereas poorly adjusted subjects are inconsistent, scoring sometimes positively, sometimes negatively, so as to produce a total result in the neighbourhood of chance. It should be mentioned that recent work by Fraser Nicol and Humphrey (20) does not entirely support this generalization.

(g) The relation of E.S.P. ability to psychotic and neurotic syndromes.

A fair amount of clinical material is available that suggests that instances of E.S.P. occur spontaneously during the course and treatment of mental disorders. Some psychiatrists have gone so far as to suggest that Psi performs a protective and dynamic function for the personality. Although a study of the evidence bearing on this possibility is of great interest, it would take us too far from our main concern, and so will not be undertaken here.

Spontaneous cases of E.S.P. occurring during analysis have been reported by many psychiatrists, ranging from Freud (22) to Bendit (6) Erenwald (17) and Sechehaye (72) among others. Reports of such incidents have aroused considerable controversy among the psychiatrists themselves, (16) and appear to activate deep-lying insecurities in many analysts. The resultant discussions tend however, to advance one's knowledge of the personalities of the analysts rather than the subject of E.S.P.

Scattered through the psychiatric literature, (c.f. 72) there are hints that suggest that patients reaching the borderline of a psychosis show a sudden increase and uprush of E.S.P. effects, which disappear if the patient gets well or becomes definitely psychotic. It is very interesting that

mystical states seem to be accompanied by similar phenomena.

If we turn from clinical observations, we find few statistical studies of the relation of mental illness to E.S.P. ability, though some material is still awaiting publication. The available material is best summarised by Rhine (55).

As has been indicated, Bates and Newton (4) found no inherent scoring differences for the different mental syndromes.

Price showed that mental illness did not abolish E.S.P. ability, and that good scores could still be obtained from such subjects.

Schulman (75) using a screened touch matching technique, grouped patients into 14 clinical classifications, and found that his Manic-depressive-depressed subjects gave **significant** positive scores, whereas those with Involutional Melancholia gave consistent, though not significant, negative scores.

An analysis by Mrs. Forster (at present awaiting publication) of all the available material on psychotic and neurotic patients, tends to confirm that Manic-depressive-depressed patients score positively, and Involutionals negatively. There is also a suggestion that Schizophrenics (excluding Paranoids) score positively.

However, varying standards of diagnosis make it hazardous to regard a diagnosis of Schizophrenia made in one institution as equivalent to that made in another, so that it is difficult to draw valid conclusions from the data. It would seem safe to say that E.S.P. ability has been found in even severely mal-adjusted patients who would co-operate to the extent of taking the tests, and that such patients may show the same differences in their scoring patterns as do the more normal personality types that have been investigated.

(h) Conclusions:-

It will therefore be seen that a considerable amount of material is available on personality differences and their relation to E.S.P. Moreover, this material hangs together and makes sense. Looked at in one way, sheep-goat attitudes are an indication of how well the subjects are adjusted to the test situation. Intro-extraversion ratings show how well they are adjusted to others. The Rosenzweig ratings also indicate the extent to which the subject is prepared to be aggressive towards other people. The Expansion-compression rating indicates whether the subject is reaching out and making contact with the environment, or whether he is shutting himself off from it. The same measure is provided by the Rorschach. Thus all these measures have the common element of contact or withdrawal as their basis.

The fact that not all investigators have been able to obtain the same results as the original workers need not be cause for too much concern, if the various difficulties which have been outlined are born in mind. It is fair to say that the originators of the techniques have had more success with them than have subsequent investigators. In parapsychological experimentation it is becoming increasingly clear that the experimenter is as important as the subjects in producing good results. (42, 78, 100-). Yet very little is known specifically about this factor. It is of course well known that certain experimenters in parapsychology never get extra-chance results, whereas others have the "Hidas Touch". Another hindrance that stands in the way of consistent results is the present state of psychological ignorance of the basic factors of personality. Until psychologists have succeeded in isolating such basic factors, if such there be, a certain amount of confusion and disagreement must be expected.

As will be shown in the next section, even the partial measure of agreement that has been obtained provides a useful basis from which to make predictions as to what effects should be found in subsequent research data, if it is genuinely the result of E.S.P.

(3) Predictions as to what trends will be shown by the Rhodes data if it is being produced by E.S.P. ability on the part of the subjects.

We are now in a position to predict what trends the Rhodes data is likely to show if it has in fact been produced by E.S.P. ability on the part of the subjects, and not merely by uncontrolled chance factors. We will later examine the data to find out if it bears out these predictions....

Prediction 1: If the subjects are divided into a high-scoring and a low-scoring group, the high-scoring subjects will be found to be more extraverted than the low-scorers.

Prediction 2: The high-scoring subjects will be found to be better adjusted than the low-scoring group.

Prediction 3: The subjects who can be classified as "sheep" in the Rhodes experiment should score more hits than the "goats", though this difference can be expected to be relatively small.

Prediction 4: If "sheep" and "goats" are further classified as "well-adjusted" and "poorly-adjusted", the "well-adjusted sheep" will score more hits than the "well-adjusted goats", and

Prediction 5: The differences in scoring rate between the "well-adjusted sheep" and "well-adjusted goats" will be bigger than for the sheep and goat groups as a whole.

Having made these predictions, we can now turn to a consideration of the personality measures used in the Rhodes experiment, and the data yielded by them.

(4) The Tests used in the Rhodes Investigation:

It was decided to use a battery of four personality tests, and a psychic questionnaire. The tests employed were the Bernreuter Personality Inventory, the Pressey X-0 Tests, the Washburne Social Adjustment Inventory, (Thaspic Edition), and the Allport-Vernon Study of Values.

(a) The Bernreuter Personality Inventory. (7)

This test measures six traits of personality adjustment:

Neurotic Tendency (scale B1-N): Persons getting high scores on this scale tend to be emotionally unstable; Self-sufficiency (scale B2-S): Persons scoring high on this scale tend to be self-sufficient; Introversion-Extraversion (scale B3-I): high scores here show introversion; Dominance-submission (scale B4-D) persons scoring high tend to dominate others in face-to-face situations; Self-confidence (scale F1-C): Persons scoring high on this scale tend to be hamperingly self-conscious and have feelings of inferiority, and finally Sociability, (scale F2-S): Persons scoring high on this scale tend to be non-social, solitary, and independent.

The reliability of these scales ranges from .85 to .91 for the B scales, and .86 to .78 for the F scales, for University students. Mosier (12) considers these reliabilities to be rather low if individual scores are to be considered. They should be adequate for the group comparisons we plan to make.

Since the Bernreuter was constructed from Thurstone's Personality Schedule, Bernreuter's Self-sufficiency Scale, Laird's Introversion Test, and Allport's A-S Reaction Study (2), its validity is thus the same as these parent tests.

As the six scores given by the Bernreuter are obtained from the same set of answers by weighting them differently, the scales show big intercorrelations. There is no reason to expect the qualities being measured to be related on psychological grounds, so that it would appear that the correlations are caused by the method of weighting the answers, and not because of any real interconnections between the traits themselves. Thus the correlations show the Introversion scale to be the same as

the Neurotic Tendency Scale. The two F scales together give all the information contained in the B scales.

Factor analysis has shown that at least six components are contributing to the B1-N scale, and hence also to the B3-I scale for Introversion, which is correlated .95 with it. It will be remembered that Humphrey found a relationship between the Introversion scores given by this scale, and E.S.P. scoring. B3-I is therefore of great interest in the present investigation, but its real nature will be impossible to interpret because of its many component factors. Further disadvantages of the Bernreuter are that hand scoring of the test is extremely laborious, and both plus and minus totals are obtained. To eliminate minus scores 200 has been added to each score in this investigation. In spite of these disadvantages the Bernreuter promised to differentiate between high and low scoring E.S.P. subjects, and thus was included in the battery.

(b) The Pressey X-O Tests: (50)

This test consists of four sub-tests. The first is intended to diagnose abnormal fears, disgusts, sex tendencies, and self-regard. The subject is asked to cross out all words in a list which are unpleasant to him. To trap irresponsible subjects, "Jokers" are included in the lists which **very seldom** arouse dislike. If the subject crosses out many of these it indicates that he is not taking the test seriously. The second sub-test is a variant of a word list association test, the subject having to cross out any words connected in his mind with the word given at the beginning of each line. The third sub-test can best be described as an ethical discrimination test, requiring the subject to cross out any words that he considers represent things that are "wrong". The final sub-test is based on Woodworth's Psychoneurotic Inventory, the subject being required to cross out all words representing things that he has worried over. This sub-test attempts to measure the relative strengths of five selected modes of reaction, and to classify the subject in terms of Paranoid, Neurotic, Self-conscious,

Melancholic, and Hypochondriacal types.

In addition Pressey makes provision for an "Affectivity or Emotionality Score", which is simply the total number of words crossed out; and an "Idiosyncrasy Score," which is the total number of words crossed out by the subject that are "modal words" --- that is, the word in each line which is the one most frequently crossed out by subjects generally.

Of his test Pressey himself says: "The scores on the entire examination are the blurred result of a number of factors, and are of relatively little importance. However, it is possible, from the mass of data yielded by the examination, to combine certain items in such a way as to obtain, from the single examination, highly differential information with reference to a number of problems....." (From Symonds, p 191, 89.)

Though Pressey himself presents the tests as research instruments, and not precise measuring devices, the writer included them in the battery for a number of reasons. He hoped that the "Affectivity Score" might give a measure of fluidity of association, rather similar to that obtained by Stuart from his Interest Inventory. It was also hoped that some correlation might be found between E.S.P. and the five clinical types diagnosed by test 4.

(c) The Allport-Vernon Study of Values: (3)

The manual supplied with this test indicates that "this study aims to measure the relative prominence of six basic interests or motives in personality: the theoretical, economic, aesthetic, social, political and religious." "The classification is based directly on Spranger's "Types of Men".

This test was validated by using the scores of groups whose characteristics were known. Thus it was found that a group of missionaries obtained the highest average

"religious value score", and a group of scientists the highest average "theoretical value score". Correlations of the scores of individual subjects with ratings by individuals gives correlation coefficients of +.45 to +.59 (When corrected for attenuation this rises to +.83) Very high correlations are not usually obtained with ratings, so these must be considered satisfactory. Reliability by the split-half method gives an r of +.72, and repeat reliabilities are +.82. These are perhaps rather low for individual work with the scores, but are satisfactory for comparison of the mean scores of the various groups. This test was included because it was hoped that the broad measures it yielded might be found to be related to E.S.P. scoring ability. In each case the higher the score, the greater the amount of that value shown by the subject.

(d) The Washburne Social Adjustment Inventory,
Thaspic Edition, (97)

According to the manual supplied with this test, "the primary purpose of the inventory is to determine the degree of social and emotional adjustment of the individual". It attempts to measure the following traits:

(1) Truthfulness: A low score indicates truthfulness, or freedom from deliberate or intentional inaccuracies in answering the questions.

(2) Happiness: Measures the degree to which the subject is contented, has a sense of well-being, and the feeling that life is worth-while. The lower the score the happier the subject.

(3) Alienation: A measure of the sense of social membership, and acceptance, also a sense of basic similarity, of common humanity with others, and of psychological security and emotional stability in social situations. The lower the score, the more identified the subject is with the group.

(4) Sympathy: a measure of the subject's sensitive, empathetic, non-negative responsiveness to people. The lower the score, the more sympathetic the subject.

(5) Purpose: a measure of the degree to which the subject's desires involve definite planning, evaluation, selection and effort. The lower the score, the more purposeful the subject.

(6) Impulse-Judgement: "This term may be defined as the ability to judge well between conflicting impulses, so that satisfactions that are recognised as greater, but more remote or more difficult, are not discounted in favour of easier or more immediate but obviously lesser satisfactions". The lower the score, the more judgement the subject exercises.

(7) Control: a measure of the sense of self-control, self-regulation, and the ability to make and execute plans. The lower the subject's score, the more control he possesses.

(8) Subtotal Score: These separate scores are added to give a "sub-total score" indicating the general social adjustment. The lower this score, the better adjusted the individual. Provision is also made for the subject to record his wishes, and if desired these can also be scored, and their contribution included in a final total. It is not proposed to include the wishes score in the present investigation.

The validity of this test was obtained by retaining only questions which were proved to discriminate between socially adjusted and socially mal-adjusted individuals.

The reliability by retest is +.92 for the whole inventory, and from +.80 to +.90 for the sub-tests. This indicates that the inventory is sufficiently reliable for individual diagnosis.

The Washburne was included in the test battery to provide a measure of personal adjustment, since Humphrey,

using the Bernreuter, and Schmeidler, using the Rorschach, had reported relationships between E.S.P. scores and the adjustment scores of the subjects. It was hoped that by using the Washburne, more precise information would be obtained as to the kind of adjustment which was the operative factor, since three of the traits in the Washburne pertain primarily to emotional adjustments to other people, whereas three of them pertain primarily to self-organisation and self-regulation.

(e) The Psychic Questionnaire:

This was a preliminary exploratory device drawn up by the writer to obtain certain information from the subjects bearing on the way they tackled the experiment, their attitudes towards it, the nature, frequency and extent of dream and other imagery which they experienced, their attitude towards, and experience of, psychical phenomena of various kinds, and various other items of information. A study of the literature of psychical research had suggested that the items included might possibly have a bearing on the possession or absence of E.S.P. ability in the subjects. Some questions were also included to test certain "hunches" of the writer as to the kind of personality type which would make a good E.S.P. subject. Thus some questions were inserted in order to gather information as to the ease or otherwise with which the subjects tended to experience states of dissociation. It was hoped that the information gathered would at any rate show which questions were worth retaining in subsequent questionnaires, and which were useless because they failed to discriminate between the subjects, or were ambiguous.

Copies of all the personality tests used, and of the Psychic Questionnaire will be found in the Appendices.

Having described the nature of the measuring instruments which were used, we can now turn to a selection of suitable statistical techniques to handle and assess the data which they produced.

(5) Statistical Methods: selection of suitable techniques:

In the analyses that follow, two main approaches have been adopted. Firstly, the subjects' personality scores have been correlated with the number of E.S.P. hits scored, using the well-known product-moment method. Although this method is popular in psychological research, it cannot be regarded as ideal, as it will only reveal if there is a linear relationship between the number of hits scored and the personality measures. Even if the correlation technique reveals no such relationship, there remains the possibility that other non-linear relationships exist, which would be equally meaningful. Suppose for instance that there exists an optimum degree of adjustment for scoring hits: that very poorly adjusted subjects fail to score hits, but so too do very well adjusted subjects. Here the graph of the relationship between the two variables is not a straight line, but a parabola. This type of relationship would not be detected by the product moment correlation technique.

A second difficulty that has to be overcome is the sheer bulk of the material to be handled. Ideally each personality rating should be correlated with each type of hit score, since there is no way of knowing which of the many types of hit will show the biggest relationships. Since a complete treatment of this sort would take a very long time, a compromise has been adopted. Exploratory correlations have been carried out between total hit score (all types of hit, all judges, uncorrected), which is the widest and least strict category of hit, and a sample of personality ratings from each test.

Thus the correlations that follow allow the biggest dilution from chance effects, and represent the least sensitive type of analysis.

The second technique which has been adopted is capable of detecting very much fainter degrees of relationship, and was suggested to the writer by Dr. J.B. Rhine. It consists of separating the subjects into a high-scoring and a low-scoring

group, and then examining the personality test scores of these two groups to find whether they show any significant differences. To assess any such differences that are observed, two methods have been used. The first of these is a comparison of the means of the two groups to find whether the differences between them are statistically significant, using the t test. The second method makes use of the Chi-square test to find whether the observed frequencies are significantly different from those to be expected on a chance basis.

The product-moment correlation method, and the mean difference method are both parametric methods of assessment, whereas the third method, based on Chi-square, is nonparametric. It has been thought advisable to use both types, as there are certain advantages and disadvantages connected with the use of each type.

Both the product-moment correlation method, and the method of mean differences have been widely used by psychologists. However Siegel, (107) points out that it has been the custom to assume that the data assessed are of such a nature as to conform to the statistical model for which the t test is valid, without troubling to test this assumption. Before any valid conclusions can be drawn as to the significance of t test results, Siegel lists the following requirements which should be fulfilled by the data:

(1) The observations must be independent. That is, the selection of any one case from the population for inclusion in the sample must not bias the chances of any other case for inclusion, and the score which is assigned to any case must not bias the score which is assigned to any other case.

(2) The observations must be drawn from normally distributed populations.

(3) These populations must have the same variance (or, in special cases, they must have a known ratio of variances).

(4) The variables involved must have been measured in at least an interval scale, so that it is possible to use the

operations of arithmetic (adding, dividing, finding means, etc.) on the scores. He defines an interval scale as one in which the distances between any two numbers on the scale are of known size, as well as bearing some relationship to each other other than mere order, as in an ordinal scale. (page 19)

Siegel continues, (page 27) "Most behavioral scientists aspire to create interval scales, and on infrequent occasions they succeed. Usually, however, what is taken for success comes because of the untested assumptions the scale maker is willing to make. One frequent assumption is that the variable being scaled is normally distributed in the individuals being tested. Having made this assumption, the scale maker manipulates the units of the scale until the assumed normal distribution is recovered from the individuals' scores. This procedure, of course, is only as good as the intuition of the investigator when he hits upon the distribution to assume.

Another assumption which is often made in order to create an apparant interval scale is the assumption that a person's answer of "yes" on any one item is exactly equivalent to his answering affirmatively on any other item. This assumption is made in order to satisfy the requirement that an interval scale have a common and constant unit of measurement. In ability or aptitude scales, the equivalent assumption is that giving the correct answer to any one item is exactly equivalent, (in amount of ability shown) to giving the correct answer to any other item."

"At the risk of being excessively repetitious, the writer wishes to emphasize here that parametric statistical tests, which use means and standard deviations (i.e. which require the operations of arithmetic on the original scores), ought not to be used with data in an ordinal scale. The properties of an ordinal scale are not isomorphic to the numerical system known as arithmetic. When only the rank of the scores is known, means and standard deviations found on the scores themselves are in error to the extent

that the successive intervals (distances between classes) on the scale are not equal. When parametric techniques of statistical inference are used with such data, any decisions about hypotheses are doubtful. Probability statements derived from the application of parametric statistical tests to ordinal data are in error to the extent that the structure of the method of collecting the data is not isomorphic to arithmetic. Inasmuch as most of the measurements made by behavioral scientists culminate in ordinal scales (this seems to be the case except in the field of psychophysics, and possibly in the use of a few carefully standardised tests), this point deserves strong emphasis." (ibid. p 26).

By contrast nonparametric tests have models which do not specify conditions about the parameters of the population from which the sample was drawn. They do rest on certain assumptions, which Siegel describes as follows: "Certain assumptions are associated with most nonparametric statistical tests, i.e., that the observations are independent and that the variable being studied has underlying continuity, but these assumptions are fewer and much weaker than those associated with parametric tests. Moreover, nonparametric tests do not require measurement so strong as that required for the parametric tests; most nonparametric tests apply to data in an ordinal scale, and some apply also to data in a nominal scale." (Ibid. p 31.)

The decision between choosing a nonparametric test or a parametric test would thus appear to be clearly in favour of the former. Unfortunately the nonparametric tests are not so powerful and sensitive as the parametric, especially where N is small, and secondly they do not use all the information contained in the data, if it reaches interval scale standards.

Consequently it has been decided to use the more sensitive parametric techniques first, even although this

may involve some untestable assumptions as to the basic nature of the data being used. If the parametric methods reveal any significant differences, it is proposed to check these particular differences using a nonparametric method.

It will be remembered that for the t test to be valid, the observations tested must be (1) independent, (2) drawn from normally distributed populations, (3) the ratios of the variances of the various populations should be known, and (4) the variables involved should be measured in at least an interval scale.

When these criteria are applied to the data which we propose to use, it is plain that (1) and (3) are fulfilled. To test (2) score distributions of the total group of subjects were drawn up for each personality score. (These are given in the Appendices). From these it will be seen that the Study of Values Test, the Bernreuter and the Pressey all produce fairly normal distributions. By contrast the Washburne Social Adjustment Inventory only produces normal distributions in the case of the Sympathy, Purpose and Subtotal scores. It is obviously no use to assess Truthfulness, Happiness, Alienation, Impulse-judgement and Control by the methods we intend to employ initially.

In two crucial cases, where the differences found approach significance (in the case of Bernreuter FIC and B3I scales) distributions for each of the groups assessed are given separately.

The remaining assumption (4) that the tests have achieved interval measurement is untestable. Moreover the respective test manuals give no information on this point. The assumption would appear very doubtful in view of the state of psychological ignorance concerning the qualities measured.

- (6) Product-moment correlations of subjects' total hit scores, (all types), with their personality test scores.

(a) Introduction:

Before embarking on this analysis, some account has to be taken of the fact that not all the subjects returned the full 50 drawings in the experiment. To allow for this, all defaulters scores have been corrected to what they would be expected to score had they returned the full 50 drawings. Thus if a subject returned 25 drawings, in which he scored 4 hits, he would be expected to score 8 hits had he returned the full 50 drawings. This score of 8 is used for correlation purposes.

(b) Statistical Method:

In the examples which follow, the formula for finding a product moment correlation r has been taken from Garrett (103 page 275). The formula for finding the standard error of the correlation has been taken from Milton Smith (104, page 76).

As an example of how the correlations are carried out let us consider calculation C45 in detail. We are correlating the Theoretical Score from the Study of Values Test, with the total hits scored on the Experimental Series of Originals by subjects from both Experimental and Control Groups. For the purpose of the calculation, let us call the Theoretical Value scores X 's, and the Total Hit Scores Y 's.

We begin by finding for each subject the square of his Theoretical Value Score, X^2 ; the square of his Total Hit Score Y^2 ; and the product of these two scores, XY ,

We then find the sum of the X 's for all the subjects, and the sum of the Y 's; and the sum of the products XY . We also need to know the total number of subjects in the group, N . These preliminary quantities were obtained by means

of a Munroe-Matic calculating machine, which does all these operations simultaneously, and accumulates the totals. All the operator has to do is to set the two scores in the machine, X and Y, press the requisite buttons, and the machine gives X, Y, X², Y² and 2XY. The totals for each of these headings are accumulated as each set of scores is fed into the machine. By checking the sum of the X's and the Y's, X², Y² and 2XY are automatically checked as well, and any errors on the part of the operator are easily discovered. Consequently the quantities given at the beginning of each calculation can be taken as correct, and an example of how they were obtained will not be given.

These values are now substituted in the formula

$$r_{XY} = \frac{N \Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[N \Sigma X^2 - (\Sigma X)^2][N \Sigma Y^2 - (\Sigma Y)^2]}}$$

giving

$$r_{XY} = \frac{(353 \times 87,346) - (11,056 \times 2,787)}{\sqrt{[353 \times 365,308 - (11,056)^2][353 \times 24,871 - 2,787^2]}}$$

Taking the top line first, we multiply N by ΣXY giving $353 \times 87,346 = 30,833,138$. Next we subtract from this the product of ΣX and $\Sigma Y = 11,056 \times 2,787 = 30,813,072$, leaving a remainder in the top line of + 20,066

Turning to the bottom line, we first multiply N by ΣX^2 giving $353 \times 365,308 = 128,953,724$. From this we take the square of $\Sigma X = 11,056^2 = 122,235,136$, giving a remainder of 6,718,588.

The same operations are performed on the bracket dealing with the Y scores, giving a remainder here of 1,012,094.

Next we multiply our two remainders together, $6,718,588 \times 1,012,094 = 6,799,842,603,272$. Taking the square root of this product gives 2,607,651. Dividing this into the top line of the formula gives $\frac{20,066}{2,607,651} = .007695$. This figure is our correlation coefficient.

To assess whether this correlation is large enough to be regarded as significant, it is necessary to find the critical ratio $\frac{r}{\sigma_{r_0}}$ (Milton Smith, 104, page 76). The formula for the standard error of the correlation is

$\sigma_{r_0} = \frac{1}{\sqrt{N-1}}$ Substituting the values we have found in this formula gives:

$$\begin{aligned}\sigma_{r_0} &= \frac{1}{\sqrt{353-1}} = \frac{1}{\sqrt{352}} = \frac{1}{\sqrt{18.76}} \\ &= .0533\end{aligned}$$

We next find the critical ratio by dividing the correlation by this quantity we have just found, σ_{r_0} giving

$\frac{.007695}{.0533}$ which gives .14 as our value for t.

Entering appropriate tables, (such as Milton Smith, 104, page 59) with this value of t shows us that the probability (chances in 100) that a value of t as large as the obtained value, or larger, could occur on the basis of chance variations in sampling (Large samples from the same population) is between 100 and 84. Thus the correlation we have found is clearly not significant.

(c) Calculations C45 to C54 follow overleaf:

Calculation
Serial No.: C 45.

Source Data
Serial No.: D 13.

Product-moment Correlation between Study of Values Theoretical Score, and "All Hits Combined" on Experimental Series Originals made by Control and Experimental Group Subjects.
(Raw scores, all judges: expectancies on the basis of 50 drawings per subject).

Number of Cases (N) = 353.

Totals:	$\sum X$	$\sum Y$	$\sum X^2$	$\sum XY$	$\sum Y^2$
	11,056	2,787	365,308	87,346	24,871

$$r_{XY} = \frac{N \cdot \sum XY - (\sum X) \cdot (\sum Y)}{\sqrt{[N \cdot \sum X^2 - (\sum X)^2] \cdot [N \cdot \sum Y^2 - (\sum Y)^2]}}$$

$$= \frac{(353 \times 87,346) - (11,056 \times 2,787)}{\sqrt{(353 \times 365,308 - 11,056^2) \cdot (353 \times 24,871 - 2,787^2)}}$$

$$= \frac{30,833,138 - 30,813,072}{\sqrt{(128,953,724 - 122,235,136) \cdot (8,779,463 - 7,767,369)}}$$

$$= \frac{+ 20,066}{\sqrt{(6,718,588)(1,012,094)}} = \frac{+ 20,066}{\sqrt{6,799,842,603,272}}$$

$$= \frac{+20,066}{2,607,651} = + .007695$$

$$\sigma_{r_0} = \frac{1}{\sqrt{N - 1}} = \frac{1}{\sqrt{353 - 1}} = \frac{1}{\sqrt{352}}$$

$$= \frac{1}{18.76} = .0533 \longrightarrow$$

$$\text{Critical Ratio: } \frac{r}{\sigma_{r_0}} = \frac{.007695}{.0533} = .14 \longrightarrow$$

Hence P lies between: 1.0 and .84 \longrightarrow

Calculation
checked ? Yes.

TABLE V.

CORRELATIONS OBTAINED BETWEEN SUBJECTS' HIT SCORES
AND PERSONALITY TEST SCORES.

(Total hits: all types, all judges, uncorrected.)
(Defaulters' hit scores corrected to the expectancy
for a full 50 drawings.)

Calculations C45 to C54.

Personality measures:-	Correlation "r"	σ_{r_0}	$\frac{r}{\sigma_{r_0}}$	P* lies between:
<u>Study of Values:</u>				
Theoretical	+.0077	.0533	.14	1.0 and .84
Economic	+.0595	.0533	1.11	.32 .23
Aesthetic	+.0217	.0533	.41	.69 .55
Social	+.0113	.0533	.21	.84 .69
Political	-.0404	.0533	.76	.55 .42
Religious	-.0410	.0533	.77	.55 .42
<u>Pressey X-C Test</u>				
Total words crossed out	-.0254	.053	.48	.69 .55
<u>Washburne Social Adjustment Invent.</u>				
Subtotal	-.036	.053	.68	.55 .42
<u>Bernreuter Personality Inv.</u>				
Confidence F1C.	-.0803	.0536	1.50	.16 .11
Sociability F2S	-.0859	.0536	1.60	exactly .11

*Probabilities have been obtained from Milton Smith: "A simplified Guide to Statistics" Table VIII, but presentation has been slightly changed to conform to the conventional method of presenting P values. Here P = .01 means that there is 1 chance in 100 that a value of t as large as the obtained value, or larger, could occur on the basis of chance variations in sampling. (Large samples from the same population.)

Transcription checked ? Yes.

(d) Conclusions drawn from Table V:

Calculations C45 to C54:

As all the above correlations are very small, and as none reach the .01 level of significance, it appears that the "Total Hit Score" is too insensitive a measure of the subjects' telepathic ability to yield any valid relationships.

However the above table suggests that the most promising measures are provided by the Bernreuter Personality Inventory scores, with the Economic Value score from the Study of Values Test the next best.

Since the correlation technique used in the above analysis is a relatively insensitive tool, and since all types of hit were used, making possible large chance dilutions, the analysis can be regarded as the least hopeful and sensitive that could be made. For any relationships to have been revealed, they would have had to have been quite marked.

The next step is to carry out an analysis on the purest measures of E.S.P. ability that the experiment has yielded, namely, the type A hits, using a more sensitive statistical procedure. If this analysis also fails to show any significant relationships, we can safely conclude that it would be fruitless to carry out any further analyses on the data.

THIRD MAIN ANALYSIS: PERSONALITY
(continued).

(7) Examination of the Mean Personality Score Differences
shown by the various groups of subjects:

(a) Introduction:

For this analysis the subjects have been divided into a number of groups, as follows:

- (a) Scoring 2 or more A hits. (d) Scoring 0 or 1 A hit.
(b) Scoring 1, 2 or more A hits. (e) Scoring 0 A hits.
(c) Scoring 1 A hit.

Since we are going to assume that group (a) contains the highest proportion of successful users of E.S.P., and group (e) the least, with the other groups intermediate between them, we first need to consider how justified such assumptions are.

Although there is a fair possibility that a subject will score an A hit by chance, the possibility of 2 A hits being scored from chance is very small. Consequently it seems safe to assume that the subjects scoring two or more A hits represent individuals successfully using E.S.P. This group represents the greatest concentration of successful E.S.P. subjects it is practicable to obtain from the data.

The group scoring 1 A hit is likely to be made up of subjects successfully using E.S.P. to score their hit, and partly by subjects who have produced an A hit from factors other than the genuine use of E.S.P. Thus it would seem safe to regard this group as representing a less concentrated sample of successful E.S.P. users than the first group.

Finally, it would seem justifiable to suppose that the least successful subjects will be found in the group that failed to score A hits. If the groups are listed in order of concentration of successful E.S.P. subjects they would come in the order given above from (a) to (e).

If the successful exercise of E.S.P. ability is related to personality factors, one would expect to obtain the most significant differences between the group scoring no A hits, and the group scoring two or more A hits, since these

give the clearest differentiation between successful and non-successful subjects.

If the subjects scoring one A hit are combined with those scoring two or more A hits, the effect will be to dilute the concentration of successful E.S.P. subjects in the group, and introduce some subjects who have scored their hit from chance. Logically one would expect that the mean differences between group (b) and group (e) would be less than the mean differences between group (a) and group (e).

Bearing this prediction in mind, we can now examine the significance of the differences in means shown by the groups.

(b) Statistical Method:

As an example of the method of assessment let us take calculation C55, dealing with the Washburne sub-total scores of the group of subjects scoring 1, 2 or more A hits, and the group scoring no A hits. For the purposes of the calculation, let us call the first of these groups (a), and the second (b). We begin by finding the total number of subjects in group (a) calling this N. Here $N = 103$.

Next we add up all the scores of these subjects, and find the sum ΣX , which in this example is 11,990. We then square each of the individual scores (the X's), and find the sum of these squared scores: $\Sigma X^2 = 1,570,932$.

We next find the mean of group (a) by dividing the sum of the scores we have found by the number of subjects in the group: $\frac{\Sigma X}{N} = \frac{11,990}{103} = 116.41$

The next step is to find the σ of the scores, by applying the formula $\sigma_X = \sqrt{M_X^2 - (M_X)^2}$

To get M_X^2 we divide the quantity ΣX^2 which we have already found by N, giving $\frac{1,570,932}{103} = 15,251.77$

We have already found $M_X = 116.41$, so we square this to get $(M_X)^2$ giving 13,551.29

Subtracting $(M_X)^2$ from M_{X^2} gives $15,251.77 - 13,551.29$ which equals $1,700.48$

To find the standard deviation we take the square root of this quantity, giving $\sigma_X = 41.24$. (This is known as Hull's method of finding a standard deviation.) To simplify the calculation, the mean is assumed to be zero, and a correction made for the error thus introduced in the formula. The formula here employed can be derived from that given by Garrett, (103, page 49-50.).

Next we have to find the σ of the mean of the scores. The formula for this is given differently by different statistical authorities, as $\sigma_M = \frac{\sigma}{\sqrt{N}}$ or as $\sigma_M = \frac{\sigma}{\sqrt{N-1}}$ so that we have to decide which of these to use on our data.

Moroney, (105, page 137) gives the first without any qualification. Garrett, (103, page 201) also gives the first version, but in a footnote comments that when the number of cases is small, below 30, $\frac{\sigma}{\sqrt{N-1}}$ gives a more accurate value. Fisher, (102, page 115), uses the first for large samples. Milton Smith, (104, page 54), points out, however, that even although \sqrt{N} is used for large samples, its use is only correct when σ is the standard deviation of the entire population, rather than the standard deviation of the sample only, as is usually the case. He goes on to show that mathematically $\sqrt{N-1}$ is the correct form for large and small samples.

The form $\sqrt{N-1}$ will tend to make the null hypothesis more difficult to reject, and will thus give a conservative assessment of any score differences significances. It will be seen too, that as N increases, the influence of the correction becomes less. It therefore appears safe to use $\sigma_M = \frac{\sigma}{\sqrt{N-1}}$ and this has been done in the calculations which follow.

Substituting the values we have found in this formula gives $\frac{41.24}{102} = \frac{41.24}{10.10} = 4.08$ for σ_{M_X}

The same procedure is followed for group (b), these scores being called the Y's. For this group $N = 252$,

$$M_Y = 114.12, \quad \sigma_Y = 32.86 \quad \text{and} \quad \sigma_{MY} = 2.07$$

Next we subtract the smaller mean from the larger. In this case the mean of group (b) is the smaller, so we get $116.41 - 114.12 = 2.29$ as the difference between the means of the two groups.

Before we can tell whether this difference is large enough to be significant, we need to calculate the σ of the difference in means. Milton Smith (104, page 57) gives the formula for this as

$$\sigma_{\text{diff}} = \sqrt{(\sigma_{MX})^2 + (\sigma_{MY})^2}$$

This formula is for two large uncorrelated samples, where $N_1 + N_2$ is more than 30 cases. The samples we are assessing conform to this model. Substituting the values we have already found in this formula gives:

$$\begin{aligned} \sigma_{\text{diff}} &= \sqrt{4.08^2 + 2.07^2} = \sqrt{16.65 + 4.28} \\ &= \sqrt{20.93} = 4.575 \end{aligned}$$

The final step is to work out the critical ratio, t :

$$t = \frac{\text{Diff. in means}}{\sigma_{\text{diff}}} = \frac{2.29}{4.57} = .50$$

Entering appropriate tables of t with this value of .50 (Such as Milton Smith, 104 page 59) tells us that there are between 55 and 69 chances in 100 that a difference in means of this size could be produced by chance. Hence the difference is clearly not significant.

This procedure is repeated for various other types of personality scores in the calculations which follow:

(c) Calculations C55 to C72:

Calculation
Serial No.: C 55.

Source Data
Serial No.: D 14.

Type of Score: Washburne Social Adjustment Inventory Subtotal.

Assessment of the significance of the difference in mean scores of the two groups shown below:

Group (a) Scoring 1, 2 or 3
A hits, raw scores, all judges:

$$\begin{array}{r} N \quad \sum X \quad \sum X^2 \\ 103 \quad 11,990 \quad 1,570,932 \\ M_X = \frac{11,990}{103} = 116.41 \end{array}$$

$$\begin{aligned} \sigma_X &= \sqrt{M_X^2 - (M_X)^2} \\ &= \sqrt{\frac{1,570,932}{103} - 116.41^2} \\ &= \sqrt{15,251.77 - 13,551.29} \\ &= \sqrt{1,700.48} = 41.24 \end{aligned}$$

$$\begin{aligned} \sigma_{M_X} &= \frac{\sigma_X}{\sqrt{N-1}} \\ &= \frac{41.24}{\sqrt{102}} \\ &= \frac{41.24}{10.10} = 4.08 \end{aligned}$$

Group (b) Scoring no A Hits:

$$\begin{array}{r} N \quad \sum Y \quad \sum Y^2 \\ 252 \quad 28,757 \quad 3,553,977 \\ M_Y = \frac{28,757}{252} = 114.12 \end{array}$$

$$\begin{aligned} \sigma_Y &= \sqrt{M_Y^2 - (M_Y)^2} \\ &= \sqrt{\frac{3,553,977}{252} - 114.12^2} \\ &= \sqrt{14,103.08 - 13,023.37} \\ &= \sqrt{1,079.71} = 32.86 \end{aligned}$$

$$\begin{aligned} \sigma_{M_Y} &= \frac{\sigma_Y}{\sqrt{N-1}} \\ &= \frac{32.86}{\sqrt{251}} \\ &= \frac{32.86}{15.84} = 2.07 \end{aligned}$$

Greater mean: 116.41
Lesser mean: 114.12
Difference: 2.29

$$\begin{aligned} \sigma_{\text{diff.}} &= \sqrt{\sigma_{M_X}^2 + \sigma_{M_Y}^2} = \sqrt{4.08^2 + 2.07^2} \\ &= \sqrt{16.65 + 4.28} = \sqrt{20.93} = 4.575 \end{aligned}$$

$$\text{Critical Ratio:} = \frac{\text{Diff. in Means}}{\sigma_{\text{diff.}}} = \frac{2.29}{4.57} = .50 \rightarrow$$

Hence P lies between .69 and .55

Calculation
checked ? Yes.

TABLE VI.

SHOWING PERSONALITY TEST SCORE MEANS FOR
 VARIOUS SUBJECT GROUPS CLASSIFIED IN TERMS
 OF UNCORRECTED A HITS SCORED.

(Calculations C55 to C72).

Personality measure:	Numbers of A hits scored by the group:				
	0	0 & 1	1	1,2 or more	2 or more
<u>Washburne:</u>					
Subtotal	114.12	-	-	116.41	108.92
<u>Pressey X-0 Test:</u>					
Melancholic:	8.95	-	-	9.27	8.38
Total crossed out:	167.08	-	-	168.34	157.96
<u>Bernreuter:</u>					
F1-C	213.16	214.88	220.23	210.83	179.917
F2-S	194.79	-	-	190.51	184.417
B3-I	197.27	-	194.96	188.699	168.08

Transcription
 checked ? Yes.

(d) Conclusions drawn from Table VI presenting personality score mean differences.

For the moment let us consider trends shown by the data without considering whether they are sufficiently marked to be regarded as statistically significant. In particular we are going to compare the means of the group scoring 0 A hits with the means of the group scoring 2 or more A hits, because the difference in E.S.P. ability shown is presumably greatest between these two groups.

It will be seen that the successful subjects are better adjusted, (as shown by the Washburne subtotal, Pressey Melancholic score, and the Bernreuter B3-I and F1-C scales, since these two latter scales are correlated +.95 with the B1-N scale for Neurotic Tendency), than the unsuccessful ones. This is consistent with the findings of Humphrey & Schmeidler on adjustment.

The successful subjects also show less affectability, (as measured by the Pressey score for the total number of words crossed out), than the unsuccessful ones. Since each word crossed out represents an emotional dislike or association, the Pressey Total can be regarded as giving a rating roughly akin to the Stuart Interest Inventory Rating. He found that his affectable subjects tended to score negatively because they tended to displace their hits, whereas his midrange group scored positively. Since the Rhodes data was scored so as to allow hits no matter how far they were displaced precognitively or postcognitively, a direct comparison of the findings of the two researches is not strictly possible. However, one would conclude from the Rhodes data that if a procedure is used that accepts a hit no matter how far it is displaced, then the subjects who react emotionally do worse than the subjects who keep the even tenor of their way, and do not become so emotionally bound up with life:: who form and react to fewer emotional associations. This is in accord with Schmeidler's findings, and extends the

findings of Stuart. (70).

The F2-S Bernreuter scale, (correlating only $+0.39$ with B3-I and $+0.11$ with F1-C), which measures sociability, shows the successful subjects to be more sociable than the unsuccessful group.

Finally the B3-I (Introversion) scale shows that the successful subjects are more extraverted than the unsuccessful group, a relationship already found by Humphrey (37). It will be remembered that it was predicted that this relationship would be found if the Rhodes data was being produced by E.S.P. ability on the part of the subjects.

Thus all the results so far discussed are completely consistent with expectancies in terms of previous research.

Next let us turn to a consideration of the scores of the group of subjects scoring only 1 A hit. It will be remembered that some of these hits will be due to chance factors, hence the introduction of such subjects may introduce chance dilutions which will blur the distinctions already found between the other two groups. On theoretical grounds we might expect the means of the groups containing subjects scoring 1 A hit to fall between the means of the group scoring 0 A hits, and the group scoring 2 or more A hits. It will be seen from the table that this has not occurred. In some cases the group scoring 1, 2 or more A hits score more than the group scoring 0 A hits. It will be seen, however, that the differences between these latter groups are smaller than the differences between the group scoring 2 or more A hits, and the group scoring 1, 2 or more A hits. Thus the group scoring 1 A hit is closer to the unsuccessful subjects than it is to the highly successful subjects.

Whereas the hierarchical order for all the groups is thus not completely consistent with that expected on theoretical grounds, yet the crucial relationships remain unaltered. The 2 or more A hit group return the smallest scores of all the groups for each of the personality measures

here investigated.

Our next concern will be to find how many of these observed differences are large enough to be regarded as statistically significant. This data is presented in Tables VII and VIII given overleaf.

TABLE VII.

SUMMARY OF THE SIGNIFICANCE OF THE DIFFERENCES
IN PERSONALITY SCORE MEANS FOR THE GROUPS OF
SUBJECTS INDICATED BELOW:

(Calculations C55 to C68)

Type A hits: all judges, uncorrected, on Experimental Series Originals only --- made by Control and Experimental Group S's.

Personality measures:	Difference between: (a) Group scoring 1, 2 or more A Hits and (b) Group scoring no A hits.			Difference between: (c) Group scoring 2 or more A hits and (d) Group scoring no A hits.		
	Mean diff:	$\frac{D}{\sigma_{diff}}$	P lies between	Mean diff	$\frac{D}{\sigma_{diff}}$	P lies between
<u>Washburne Social Adjustment Inven.</u> Subtotal:	2.29	.50	.69 .55	5.20	.67	.55 .42
<u>Pressey X-0 Test:</u>						
Melancholic Score:	.32	.58	.69 .55	.57	.66	.55 .42
Total words crossed out:	1.26	.22	.84 .69	9.12	.90	.42 .32
<u>Bernreuter:</u>						
FLC Confidence:	2.33	.22	.84 .69	33.24	1.97	.05 .045
F2S Sociability:	4.28	.65	.55 .42	10.37	.84	.42 .32
B3I Introversiion:	8.57	1.54	.16 .11	29.19	3.37	.001 .0007
			(e) Group scoring 1 A hit, and (f) Group scoring 0 or 1 A hit.			(e) Group scoring 1 A hit, and (c) Group scoring 2 or more A hits.
<u>Bernreuter:</u>						
FLC Confidence:	5.35	.45	.69 .55	40.31	2.09	.045 .036

Notes: The above values for P have been obtained from Milton Smith (104, page 59). Here a value of $P = .01$ means that there is only 1 chance in 100 that a value of t as large as the obtained value, or larger, could occur on the basis of chance variations in sampling. (Large samples from the same population.). (P values are for a two-tailed test of significance).

Transcription checked ? Yes.

(e) Conclusions drawn from Table VII on the significance of the mean personality test score differences:

From the above table it will be seen that three differences exist which are significant at the .05 level or better. *

Comparing first the groups scoring 0 A hits and 2 or more A hits, it will be seen that the hit scoring group are significantly more confident, (Bernreuter FLC scale), than the non-hit scoring group. When the group scoring 2 or more A hits is compared with the group scoring 1 A hit, the same significant difference is found: the hit scoring subjects are again significantly more confident. These differences reach the .05-.045 and .045-.036 levels of significance respectively.

Turning to the B3I Introversion score, it will be seen that the group scoring 2 or more A hits are significantly more extraverted than the group scoring 0 A hits. This is a highly significant difference, P lying between .001 and .0007.

In view of this promising result, it was considered advisable to determine the B3I score differences for other possible groups in terms of hits scored, to discover more precisely how Introversion was related to the hit scoring abilities of the groups. Further calculations were therefore made, using the same method as that used in the previous calculations. The results are presented below in Table VIII.

(f) Calculations C69 to C72:

* The P values given are for a two-tailed test of significance, since for the moment we are concerned to find if any significant differences exist, regardless of their direction. When we come to examine how far our predictions are born out by the results, it will be in order to use a one-tailed test of significance, since we have predicted the direction the difference will take. For a one tailed test the above P values should be halved, increasing the significance of the results.

TABLE VIII.

SUMMARY OF THE BERNREUTER B3I- INTROVERSION SCORES:
SIGNIFICANCE OF MEAN DIFFERENCES FOR VARIOUS GROUPS
IN TERMS OF A HITS SCORED.

(A) COMPARISON WITH GROUP SCORING 0 A HITS:
 (Calculations C 65, 66, 69)

A hits scored by group:	B3I Mean score:	Mean diff.	$\frac{D}{\sigma_{diff.}}$	P lies between:
0 A hits:	197.27			
1 A hit:	194.96	2.31	.37	.84 .69
1 or more	188.699	8.57	1.54	.16 .11
2 or more	168.08	29.19	3.37	.001 .0007

(B) COMPARISON WITH GROUP SCORING 1 A HIT:
 (Calculations C 70, 71).

1 A hit:	194.96			
1 or more	188.699	6.261	.87	.42 .32
2 or more	168.08	26.88	2.75	.007 .005

(C) COMPARISON WITH GROUP SCORING 1 OR MORE A HITS.
 (Calculation C 72).

1 or more	188.699			
2 or more	168.08	20.619	2.19	.36 .28

(D) RELATION BETWEEN NUMBER OF A HITS SCORED AND DEGREE OF EXTRAVERSION

Group scoring 0 A hits B3I mean score 197.27 (most introverted)

Group scoring 1 A hit: 194.96

Group scoring 1 or more 188.699

Group scoring 2 or more 168.08 (most extraverted)

Notes: Values for P have been taken from Milton Smith, (104, page 59), and show the probability (chances in 100) that a value of t as large as the obtained value or larger, could occur on the basis of chance variations in sampling. (Large samples from the same population). In table VII above P = .01 means one chance in 100. (P values are for a two-tailed test.)

Transcription checked ? Yes.

(g) Conclusions drawn from Table VIII: Significance of the Bernreuter B3I mean score differences for various A hit scoring groups:

It will be seen that the results shown in this table are completely consistent: there is a perfect correlation between the hierachical order in terms of hits scored, and the amount of Extraversion shown by the groups. Moreover, as the numbers of hits scored by the groups increases, so too does the significance of the differences found. This is not an artifact due to increasing numbers of subjects in these groups, for the number of subjects falls off as the number of hits increase. There can be little doubt that a real relationship exists between number of hits scored and degree of extraversion in the subjects. However this can only be taken as strengthening the E.S.P. hypothesis if it can be shown that it is not due to artificial factors.

It could be argued, for instance, that the Originals used in the experiment had been chosen in a special way, and that this way made it easier for extraverts to score hits on them in the Main Experiment than introverts. It is therefore necessary to test this hypothesis. It can be done by finding the mean Bernreuter scores of the groups of subjects who had suggested the titles of the Originals in the Preliminary Frequency Experiment, and comparing these with the mean scores of groups of subjects who had actually scored hits in the Main Experiment. The same statistical method previously employed is used again here.

(h) Calculations C73 to C76:

TABLE IX.

SIGNIFICANCE OF THE DIFFERENCES BETWEEN MEAN INTRO-
VERSION SCORES, (B3I + 200), OF GROUPS OF SUBJECTS
(a) SUGGESTING TITLES FOR CSO AND ESO ORIGINALS
IN THE PRELIMINARY FREQUENCY EXPERIMENT AND
(b) SCORING UNCORRECTED A HITS IN THE MAIN
EXPERIMENT.

(Calculations C72 to C76)

<u>Groups:</u>	Means:	Mean differences:	$\frac{D}{\sigma_{M_{diff}}}$	Chances in 100 that D is signif*
1 title:	194.96			
1 A hit:	189.42	5.54	.76	77 to 79
1 or more titles:	188.699			
1 or more A hit:	189.94	1.241	.19	56 to 58
2 or more titles:	192.33			
2 or more A hits:	168.08	24.25	1.93	97 to 98
2 or more titles	192.33			
0 A hits:	197.27	4.94	.48	67 to 69

* From Garrett: Statistics in Psychology and Education, Table 34, page 213: showing the number of chances in 100 that the true difference is greater than zero. (103).

Transcription checked ?
Yes.

(i) Conclusions drawn from Table IX: Significance of the differences in mean Introversion Scores for title-suggesting groups in the Preliminary Frequency Experiment, and hit-scoring groups in the Main Experiment:

It will be remembered that the hypothesis that we are testing is that Extraverts were more prone to suggest singles titles than were Introverts in the Preliminary Frequency Experiment, and thus it was easier for Extraverts to score hits in the Main Experiment.

The data presented in Table IX does not support this hypothesis. It will be seen that no significant differences in mean extraversion scores exist between groups scoring 1 title and 1 A hit, or suggesting 1 or more titles and scoring 1 or more A hits, or more crucially, between groups suggesting 2 or more titles and scoring 0 A hits.

By contrast the only difference which approaches significance is that between the group suggesting 2 or more titles and that scoring 2 or more A hits. The title suggesting group are more introverted than the hit scoring group.

This is what would be expected on psychological grounds: it would seem more likely that Introverts would have unusual, individual ideas, and Extraverts would be more banal producing common and popular ideas. Thus if anything, the way in which the Originals were chosen would seem to make it easier for Introverts to score hits in the Main Experiment, than Extraverts. Since it has been found that Extraverts did better in the Main Experiment, this cannot therefore be explained as an artifact of the experimental design. Had it been found that no significant difference existed between the group scoring two or more A hits and the group suggesting two or more titles, but that a significant difference existed between the group suggesting 2 or more titles and scoring 0 A hits, then this would have been strong evidence for the anti-E.S.P. hypothesis.

The evidence thus tends to confirm that the extraversion of the subjects is linked to their E.S.P. hit-scoring

ability, and not due to experimental artifacts.

It now remains to see whether the data will yield essentially the same results when assessed by nonparametric methods.

(8) Nonparametric assessment of the Personality Test Score differences yielding significant results by parametric statistical assessment:

(a) Introduction:

It will be remembered that the Mean Difference method so far used can only give a completely accurate statement of significance if the data to which it is applied conforms to four requirements (1) the observations used should be independent, (2) they should be drawn from normally distributed populations, (3) the ratios of the variances of the different populations should be known, and (4) the variables involved should be measured in at least an interval scale.

Since it is not known whether the Bernreuter Inventory succeeds in achieving interval ^{measurement} it would seem advisable to check the results already obtained by using a method of assessment which does not depend on the above assumptions for its validity. It is proposed to use the Median Test described by Siegel in "Nonparametric Statistics", page 111, since this is based on the χ^2 procedure already used in some calculations.

(b) Statistical Method:

According to Siegel "The median test is a procedure for testing whether two independent groups differ in central tendencies. More precisely, the median test will give information as to whether it is likely that two independent groups (not necessarily of the same size) have been drawn from populations with the same median. The null hypothesis is that the two groups are from populations with the same median; the alternative hypothesis may be that the median of one population is different from that of the other (two-tailed test) or that the median of one population is higher than that of the other (one-tailed test). The test may be used

whenever the scores for the two groups are at least in an ordinal scale".

Siegel outlines the method as follows: "To perform the median test, we first determine the median score for the combined group (i.e., the median for all scores in both samples). Then we dichotomize both sets of scores at that combined median, and cast these data in a 2 X 2 table:

	Group I	Group II	Total
No. of scores above combined median	A	B	A + B
No. of scores below combined median	C	D	C + D
Total	A + C	B + D	N = n ₁ + n ₂

"Now if both group I and group II are samples from populations whose median is the same, we would expect about half of each group's scores to be above the combined median and about half to be below. That is, we would expect frequencies A and C to be about equal, and frequencies B and D to be about equal.

"It can be shown that..... if A is the number of cases in group I which fall above the combined median, and if B is the number of cases in group II which fall above the combined median, then the sampling distribution of A and B under the null hypothesis (H₀ is that A = $\frac{1}{2}n_1$ and B = $\frac{1}{2}n_2$) is the hypergeometric distribution

$$p(A,B) = \frac{\binom{A+C}{A} \binom{B+D}{B}}{\binom{n_1+n_2}{A+B}}$$

Therefore if the total number of cases in both groups (n₁ + n₂) is small, one may use the Fisher Test" (The Fisher Exact Probability Test) "to test H₀. If the total number of cases is sufficiently large, the χ^2 test with df = 1 may be used to test H₀.

"When analyzing data split at the median, the researcher should be guided by these considerations in choosing between the Fisher test and the χ^2 test:

"(1) When $n_1 + n_2$ is larger than 40, use χ^2 corrected for continuity, i.e., use the formula

$$\chi^2 = \frac{N \left(\left| AD - BC \right| - \frac{N}{2} \right)^2}{(A + B)(C + D)(A + C)(B + D)}$$

(2) When $n_1 + n_2$ is between 20 and 40 and when no cell has an expected frequency of less than 5, use χ^2 corrected for continuity (formula as above). If the smallest expected frequency is less than 5, use the Fisher test.

(3) When $n_1 + n_2$ is less than 20, use the Fisher test.

One difficulty may arise in the computation of the median test: several scores may fall right at the combined median. If this happens, the researcher has two alternatives: (a) if $n_1 + n_2$ is large, and if only a few cases fall at the combined median, these few cases may be dropped from the analysis, or (b) the groups may be dichotomised as those scores which exceed the median and those which do not. In this case, the troublesome scores would be included in the second category".

As an example of the method, let us consider calculation C77, dealing with Bernreuter FLC Confidence scores of a group scoring 2 or more A hits, and a group scoring 0 A hits. (The first group contains 24 subjects, the second 246).

We start by finding the combined number of subjects from both groups. Since $n_1 = 24$ and $n_2 = 246$, $\sum N = 270$. The median of the group thus falls between the 135th and 136th scores, i.e. between the step intervals 11 and 12. Splitting the group 12 to 289, and 11 to -185 gives the following frequencies:

	S's above mdn:	S's below mdn:	Totals:
Group scoring 0 A hits:	A 127	B 119	246
Group scoring 2 or more A hits:	C 8	D 16	24
Totals:	135	135	270

Before beginning the calculation proper, we check to see

whether the smallest expected frequency is greater than 5. Cell C appears to be the smallest, so we find the theoretically expected frequency by multiplying the marginal totals of this cell 135×24 , and dividing the result by $N = 270$. This gives 12, which is greater than 5. It is therefore in order to use the χ^2 technique. Since $n_1 + n_2$ is greater than 40, it is appropriate to use the χ^2 test corrected for continuity. This correction is given by the value $-\frac{N}{2}$ in the formula below:

$$\chi^2 = \frac{N \left(|AD - BC| - \frac{N}{2} \right)^2}{(A + B)(C + D)(A + C)(B + D)}$$

Next we find the value for the bottom line of the formula:

$$(A+B)(C+D)(A+C)(B+D) = 246 \times 24 \times 135 \times 135 = 107,600,400$$

Next we find $AD-BC = 2032-952 = 1080$ Squaring this gives 1,166,400, and multiplying by $N = 270$ gives 314,928,000

However this has not been corrected for continuity*. If we wish to correct for continuity, we proceed as follows;

First we find $\frac{1}{2}N = \frac{1}{2}$ of 270 = 135. We subtract this from

$|AD-BC|$ giving $1080 - 135 = 945$ Squaring this gives 893,025 and multiplying by N gives 241,116,750. We can now substitute the values found in the formula:

Uncorrected for continuity:

$$\chi^2 = \frac{314,928,000}{107,600,400} = 2.93$$

(1df)

P lies between .1 and .05

If it is held that the direction of the difference in the groups can be predicted, then a one tailed test of significance is appropriate, and the above probabilities can be halved, e.g.

P lies between .05 and .025

Corrected for continuity:

$$\chi^2 = \frac{241,116,750}{107,600,400} = 2.24$$

(1df)

P lies between .2 and .1

P lies between .1 and .05.

(c) Calculations C77 to C80 follow:

Table X:

*For a fuller discussion of the circumstances under which it is necessary to correct for continuity, see page 308.

Calculation
Serial No: C 77.

Source Data
Serial No.: D 15

Nonparametric Median Test.

Bernreuter Fl-C Confidence Scores for groups of subjects in terms of uncorrected A hits. (Including subjects not returning all 50 drawings, but excluding subjects returning no drawings at all.)

Combination of:

(a) Group scoring 2 or more A hits
(N = 24) with

(b) Group scoring 0 A hits (N = 246)

giving a Median Score falling between the 135th and 136th score, i.e. between the step intervals 11 and 12. Splitting the groups 12-289 and 11 to -185 gives the following frequencies:

	Scores above Median		Scores below Median		Totals
Group scoring 0 A hits:	A	127	B	119	246
Group scoring 2 or more A hits:	C	8	D	16	24
Totals:		135		135	270

Smallest theoretical frequency is $C_T = 135 \times 24 \div 270$ which is greater than 5.

$$(A+B)(C+D)(A+C)(B+D) = 246 \times 24 \times 135 \times 135 = 107,600,400$$

$$|AD - BC| = 2,032 - 952 = 1,080$$

$$1,080^2 = 1,166,400$$

$$1,080^2 \times N = 314,928,000$$

$$\frac{1}{2}N = 135 \quad |AD - BC| - \frac{1}{2}N = 1,080 - 135 = 945$$

$$945^2 = 893,025$$

$$945^2 \times N = 241,116,750$$

$$\chi^2 \text{ (uncorrected for continuity:)} \\ \frac{N(|AD - BC|)^2}{(A+B)(C+D)(A+C)(B+D)}$$

$$\frac{314,928,000}{107,600,400}$$

2.93 (for 1 d.f.)

P lies between

.1 and .05

$$\chi^2 \text{ (corrected for continuity:)} \\ \frac{N(|AD - BC| - \frac{1}{2}N)^2}{(A+B)(C+D)(A+C)(B+D)}$$

$$\frac{241,116,750}{107,600,400}$$

2.24 (for 1 d.f.)

P lies between

.2 and .1

Calculation
checked ? Yes.

TABLE X.

ASSESSMENT OF THE SIGNIFICANCE OF THE BERNREUTER FLC & B3I PERSONALITY SCORE DIFFERENCES BETWEEN (a) THE GROUP SCORING 2 OR MORE A HITS, AND (b) THE GROUP SCORING 0 A HITS, USING THE NONPARAMETRIC MEDIAN TEST.

(Calculations C77 to C80).

(A) EXCLUDING ONLY THOSE SUBJECTS RETURNING NO DRAWINGS WHATSOEVER IN THE MAIN EXPERIMENT.

	Uncorrected for continuity:	Corrected for continuity:
FLC: χ^2 (1 df.)	2.93	2.24
P for two-tailed test:	between .1 & .05	between .2 & .1
P for one-tailed test:	between .05 & .025	between .1 & .05
B3I: χ^2 (1 df.)	4.58	3.71
P for two-tailed test:	between .05 & .02	between .1 & .05
P for one-tailed test:	between .025 & .01	between .05 & .025

(B) EXCLUDING ALL SUBJECTS NOT RETURNING THE FULL 50 DRAWINGS IN THE MAIN EXPERIMENT:

	Uncorrected for continuity:	Corrected for continuity:
FLC: χ^2 (1 df.)	2.52	1.84
P for two-tailed test:	between .2 & .1	between .2 & .1
P for one-tailed test:	between .1 and .05	between .1 & .05
B3I: χ^2 (1 df.)	2.59	1.9
P for two-tailed test:	between .2 & .1	between .2 & .1
P for one-tailed test:	between .1 & .05	between .1 & .05

The P values given above have been obtained from Siegel's Table of Critical Values for Chi Square, p249 in "Nonparametric Statistics" (107).

Transcription checked ?
Yes.

(d) Conclusions drawn from Table X: Assessment of the significance of the Bernreuter FLC & B3I scores by the Nonparametric Median Test:

Table X has been set out in such a way as to reveal the influence of various factors which might possibly be affecting the results.

Firstly, P values for both one-tailed and two-tailed tests of significance are given. If it is felt that we can predict that the groups will show a difference in Bernreuter scores, but not the direction of the difference, then the two-tailed test gives the appropriate P values. However readers may feel that the work of previous investigators makes it possible to predict that the multi-hit-scoring subjects will be more extraverted than the subjects who fail to score hits, in which case the one-tailed test gives the appropriate assessment of significance. In view of the predictions made earlier, on the basis of the work of Humphrey, Schmeidler and others, the writer himself feels that the one-tailed test is the more appropriate one.

Secondly, two sets of χ^2 values have been given. One of these has been corrected for continuity, the other not. It will be seen that the corrected values give the most conservative assessment of significance.

Thirdly, the data has been broken down in terms of groups containing subjects not returning the full 50 drawings in the Main Experiment, and groups purged of all subjects failing to return the full 50 drawings. This has been done because it was felt that the more drawings returned by a subject, the better chance he had of scoring A hits. This type of conscientious subject would probably also show a different personality makeup from the subject who forgot to do some of the required drawings. It might be, therefore, that these relationships might bring about a spurious correlation of personality traits with hit scoring ability, unrelated to basic correlation of E.S.P. ability with personality makeup.

Examination of the Table will show

(1) The significances obtained from the Median Test are generally less than those obtained by the Mean difference method.

(2) Correcting for continuity lowers the significance of the differences.

(3) Eliminating subjects failing to return the full 50 drawings lowers the significance of the differences.

(4) The significances obtained range between .2 and .01. They are therefore suggestive, even though they may fail to reach the .05 level in some cases.

Though at first sight these results may cast doubt on the significant differences found by the mean difference method, there are several reasons why this nonparametric method should tend to produce lowered significances:

(1) The Median Test is admittedly less sensitive than the Mean-difference test.

(2) Significance tends to fall as the number of cases gets less. The group scoring 2 or more A hits was initially only 24 strong. The procedure in the Median Test is to consider halves of the group, thus halving the number of cases in each cell, and dropping the significance of the differences.

(3) Purging the group of subjects failing to return the full 50 drawings further lowers the number of cases, and hence the significance.

(4) Despite these factors, a number of differences verge on the .01 level of significance, and none drop lower than .2

Since the significances obtained are not high, there is necessarily some ambiguity in interpreting the data, but the writer would submit that the general trend of the results is to confirm, rather than to contradict, the previous findings from the Mean-differences assessments.

(9) Brief summary of the findings so far obtained from the three methods of assessment used on the personality test scores:

(a) No significant product-moment correlations were found between personality test scores and the subjects' scores for all types of hit combined.

(b) The clearest differences in mean personality test scores were found between the group failing to score A hits and the group scoring 2 or more A hits: that is, using the purest type of hit.

(c) The A hit scoring group were more confident (Bernreuter FLC score), less affectable (Pressey Total words crossed out), more extraverted (Bernreuter B3I score) and more adjusted (Washburne Subtotal score) than the group failing to score A hits. All these trends were predicted on the basis of the work of previous investigators. However not all of them were marked enough to reach the .05 level of significance.

(d) The FLC (confidence) score differences reach the .05 level.

(e) The B3I (introversion) scores reach the .001 level of significance.

(f) Two hypotheses alternative to the E.S.P. hypothesis were examined, but it was found that they failed to account for the relationship between extraversion and hit scoring ability.

(g) In general the median tests carried out on FLC and B3I scores produced lesser significances than those obtained from the mean difference method, but tended to support rather than contradict the previous findings.

(h) The presence of predicted relationships between personality scores and A hit scoring ability strengthens the hypothesis that such A hits were not being produced by chance alone, but that some of them were being produced by the use of genuine E.S.P. ability on the part of the subjects.

(10) Sheep-Goat Ratings: Material from the Psychic Questionnaire:

(a) Introduction:

In this section of the analysis it is proposed to examine only a small portion of the material from the Psychic Questionnaire, namely that bearing on the attitudes of the subjects to E.S.P. The other material will be assessed in a later section.

To obtain the necessary information, the subjects were asked to answer the following question in the Psychic Questionnaire:

"Indicate what your attitude to telepathy was at the beginning of the experiment by underlining (a), (b) (c) or (d) :

- (a) Sceptical
- (b) No convictions either way
- (c) Convinced that it occurs.
- (d) Convinced that you are personally telepathic."

By means of their answers to this question the subjects were separated into three groups:

The Goats: those underlining (a).... the sceptics.

The Open-minded: those underlining (b).

The Sheep: those underlining either (c) or (d):
the believers.

The numbers falling into each of these categories give an interesting indication of the opinion of university students at Rhodes towards the possibility of telepathy. The attitudes of about one third of the total student population at Rhodes were sampled, 353 subjects in all. Of these, 162 were open-minded, 134 were convinced of the existance of E.S.P., and the remaining 57 were sceptical.

It will be remembered that Schmeidler's findings were that the believing sheep returned higher scores than did the sceptical goats. We must now analyse the Rhodes data to

find whether it also shows this relationship.

(b) Statistical methods used:

It will be realised that once the subjects have been allocated to the Sheep, Goat or Open-minded groups, two main tasks confront us in assessing the data. The first of these is to present the A hit scores made by the three groups in a form in which they can be directly compared. It is necessary, for instance, to allow for the fact that not all the A hit scoring subjects returned the full 50 drawings. The second step will be to test any differences that are revealed to find whether they are large enough to be statistically significant.

The first of these requirements has been met by calculating the success per trial for each of the three attitude groups. These figures have been obtained by dividing the total number of A hits scored by the particular group by the total number of trials made by the subjects in that group. Since this merely involves listing the subjects in the three classes, and then simple transcription and addition, this part of the calculation has not been set out in detail. The results are presented graphically.

To test the significance of the differences shown, the familiar Chi-square procedure, originally used in the First Main Analysis, has again been used. Consequently it has not been thought necessary to go step by step through a sample calculation to explain the method. Briefly, the deviation of an observed frequency from a theoretically expected frequency is found, the deviation squared, and then divided by the expected frequency. This is repeated for each category in which we are interested. The separate values are totalled to give χ^2 , and the number of degrees of freedom determined. Reference to the appropriate tables, (e.g. 54, pages 369-371), then tells us whether our observed frequencies depart significantly from the theoretically expected frequencies on the hypothesis we are testing. In calculation C 81 overleaf, the observed numbers of hits obtained by Sheep, Goat and Open-

minded subjects are compared with the numbers that would be expected if each of the groups were equally good at using E.S.P.

In calculation C82 the same hypothesis is tested, but the Open-minded subjects have been eliminated, and only the Sheep and Goats used.

In private correspondence with the writer, Dr. J.G. Pratt has raised the question as to whether this method, which uses the trials as units rather than the subjects, is valid. He comments: "Your method of calculation is ok, but there is a question regarding whether this is the appropriate evaluation to use. Presumably, you are interested in testing whether there is any consistent relation among the three attitudes and rate of scoring. To do this, you must take the subject as the unit, not the individual trial. If you take the latter, there is no way to tell whether a difference is contributed by one subject or by many subjects, and only in the latter case would you be entitled to draw a conclusion regarding the relation between attitude and ESP. It seems clear that there is no significant relation to attitude, so you would not lose any "conclusion" by using the test of subject consistency, such as you have applied in Calculation No. C 83."

Upon reflection it appeared to the writer that the fact that type A hits were being used made it possible to say with fair assurance that any results obtained could not be due to large score contributions from a few individuals in the groups. It will be seen that of the subjects answering the Psychic Questionnaire, only two scored 3 A hits, twenty-two scored 2 A hits, and eighty-two scored 1 A hit. The actual distribution of subjects was as follows:

	<u>Sheep:</u>	<u>Openminded:</u>	<u>Goats;</u>
No. of subjects scoring			
0 A hits:	91	113	43
No. scoring 1 A hit:	34	37	11
No. scoring 2 A hits:	7	12	3
No. scoring 3 A hits:	2	0	0
Total subjects	134	162	57
Total hits	54	61	17

It would appear fairly safe to assume therefore, that any differences found would be due to general subject tendencies and not to a few large contributions from one or two subjects.

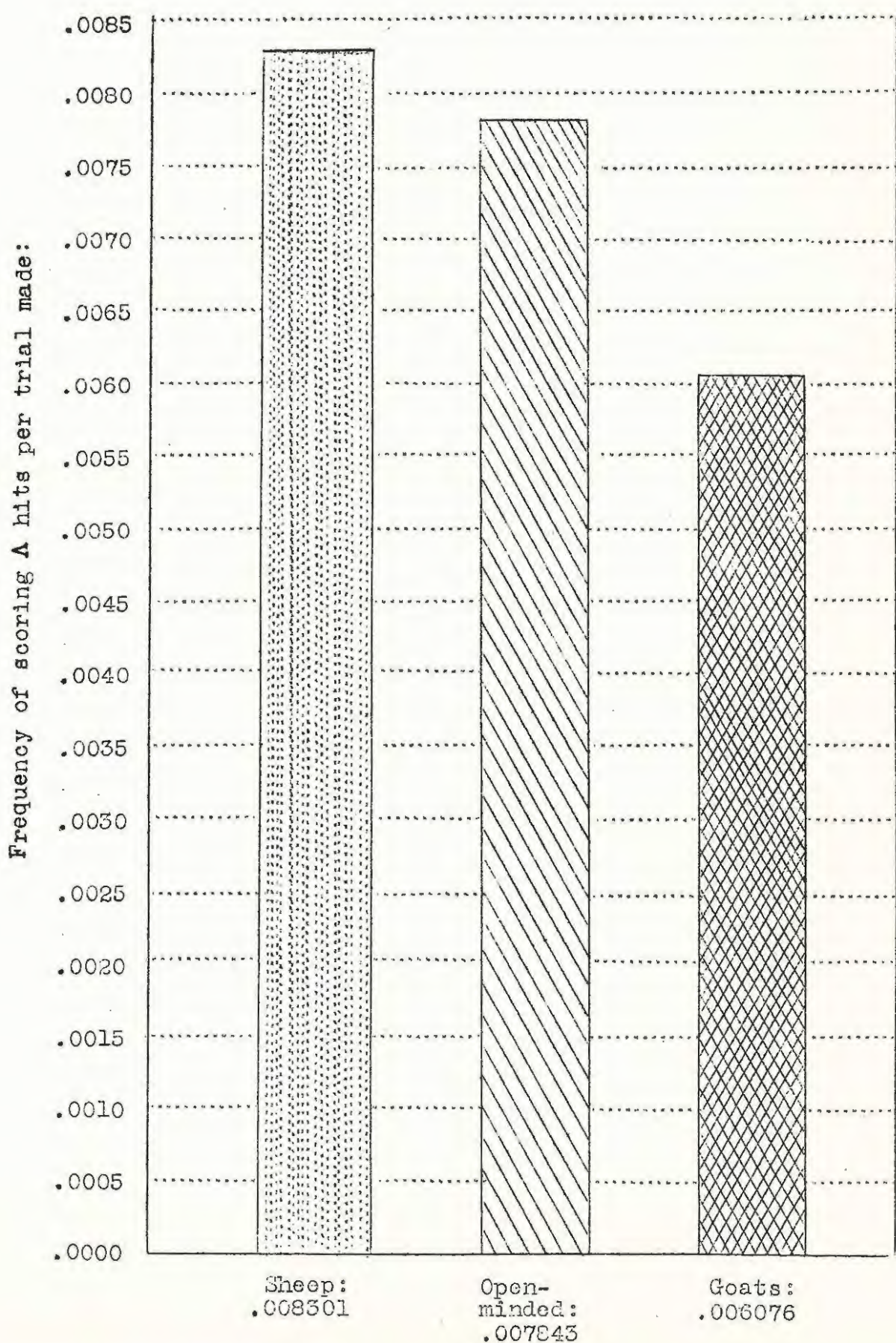
However, as an additional safeguard the method advocated by Dr. Pratt has been used in calculation C 83. In this calculation the number of subjects falling into the group sans A hits, and the group scoring 1 or more A hits for Sheep and Goats are compared with frequencies to be expected on a chance basis.

(c) Graphical presentation of data, and
calculations C 81 to C 83:

GRAPH No: 1

SHEEP-GOAT ATTITUDE ANALYSIS:

GRAPH SHOWING THE INFLUENCE OF THE SUBJECTS' ATTITUDES ON THE FREQUENCY OF SCORING A HITS.
(All judges, uncorrected.)



Calculation
Serial No: 081

Source data sheets
not reproduced.

Comparison of the number of A hits (All judges, uncorrected), scored by Sheep, Goat and Open-minded subjects with the numbers to be expected on a chance basis.

Observed frequencies:

Subjects:	Total Trials	Hits	Misses	Hit av. per trial
Goats:	2,798	17	2,781	.006076
Open-minded:	7,778	61	7,717	.007843
Sheep:	6,505	54	6,451	.008301
Totals:	17,081	132	16,949	

Expected frequencies from chance:

Goats should score $\frac{2,798}{17,081}$ of the total hits.

$$= \frac{132}{1} \times \frac{2,798}{17,081} = \frac{369,336}{17,081} = \underline{21.6226 \text{ hits.}}$$

Hence expected misses = 2,798 - 21.6226 = 2,776.3774 misses

Open-minded should score $\frac{7,778}{17,081}$ of the total hits.

$$= \frac{132}{1} \times \frac{7,778}{17,081} = \frac{1,026,696}{17,081} = \underline{60.1075 \text{ hits.}}$$

Hence misses = 7,778 - 60.1075 = 7,717.8925 misses

Sheep should score $\frac{6,505}{17,081}$ of the total hits.

$$= \frac{132}{1} \times \frac{6,505}{17,081} = \frac{858,660}{17,081} = \underline{50.2699 \text{ hits.}}$$

Hence misses = 6,505 - 50.2699 = 6,454.7301 misses

	Hits:	Misses	Totals:
Goats:	Observed: 17	2,781	
	Expected: 21.6226	2,776.3774	2,798
	Difference: 4.6226	4.6226	
Open-minded:	Observed: 61	7,717	
	Expected: 60.1075	7,717.8925	7,778
	Difference: 0.8925	0.8925	
Sheep	Observed: 54	6,451	
	Expected: 50.2699	6,454.7301	6,505
	Difference: 3.7301	3.7301	
TOTALS:	132	16,949	17,081

Calculation
serial No: C81 continued: page 2.

A Chi-square value is found for each cell from the formula:

$$\frac{(\text{Difference})^2}{\text{Expected frequency}}$$

For the Goats: = $\frac{4.6226^2}{21.6226}$ + $\frac{4.6226^2}{2,776.3774}$

= $\frac{21.3684}{21.6226}$ + $\frac{21.3684}{2,776.3774}$

= .9882 + .0077

For the Open-minded: = $\frac{.8925^2}{60.1075}$ + $\frac{.8925^2}{7,717.8925}$

= $\frac{.7965}{60.1075}$ + $\frac{.7965}{7,717.8925}$

= .0133 + .0001

For the Sheep: = $\frac{3.7301^2}{50.2699}$ + $\frac{3.7301^2}{6,454.7301}$

= $\frac{13.9156}{50.2699}$ + $\frac{13.9156}{6,454.7301}$

= .2768 + .0021

The Chi-square values for each cell are therefore as follows:

Subject category:	Hits:	Misses:	Chi square total:
Goats:	.9882	.0077	
Open-minded:	.0133	.0001	
Sheep:	.2768	.0021	
Totals:	1.2783	.0099	1.2882

Since both row and column totals are fixed, the degrees of freedom are given by the formula: $d.f. = (r - 1)(c - 1)$ where r is the number of rows, and c the number of columns.

Hence $d.f. = (3 - 1)(2 - 1) = 2 \times 1 = 2$ degrees of freedom

Entering Fisher's table of Chi-square, with $\chi^2 = 1.2882$ and two degrees of freedom we find

P lies between .70 and .50 which is not significant.

Calculation
checked ? Yes.

Calculation
Serial No: C82

Source Data
Sheets not reproduced.

Comparison of the number of A hits (all judges, uncorrected) scored by the Sheep and Goat subjects only, with the frequencies that would be expected on a chance basis. (Open-minded subjects have been excluded from this calculation.)

Observed frequencies:

Subject category:	Total trials	Hits	Misses	Hit av. per trial
Goats:	2,798	17	2,781	.006076
Sheep:	6,505	54	6,451	.008301
Totals:	9,303	71	9,232	

Expected frequencies:

$$\text{Goats: } \frac{2,798}{9,303} \times \frac{71}{1} = \frac{198,658}{9,303} = \underline{21.3542 \text{ hits.}}$$

$$\text{Hence misses} = 2,798 - 21.3542 = \underline{2,776.6458 \text{ misses.}}$$

$$\text{Sheep: } \frac{6,505}{9,303} \times \frac{71}{1} = \frac{461,855}{9,303} = \underline{49.6458 \text{ hits}}$$

$$\text{Hence misses} = 6,505 - 49.6458 = \underline{6,455.3542 \text{ misses.}}$$

Subject category:	Hits:	Misses:	Totals:
Goats:	Observed:	17	2,798
	Expected	21.3542	2,776.6458
	Difference:	4.3542	4.3542
Sheep:	Observed:	54	6,505
	Expected:	49.6458	6,455.3542
	Difference	4.3542	4.3542
Totals:	71	9,232	9,303

A χ^2 value is found for each cell of this table from the formula:-

$$\frac{\text{Difference}^2}{\text{Expected frequency.}}$$

Calculation

Serial No: 082 continued: page 2.

For the Goats:

Hits:

Misses:

$$\frac{4.3542^2}{21.3542}$$

$$\frac{4.3542^2}{2,776.6458}$$

$$\frac{18.9591}{21.3542}$$

$$\frac{18.9591}{2,776.6458}$$

.8878

.0068

For the Sheep:

$$\frac{4.3542^2}{49.6458}$$

$$\frac{4.3542^2}{6,445.3542}$$

$$\frac{18.9591}{49.6458}$$

$$\frac{18.9591}{6,445.3542}$$

.3819

.0029

Summing these values gives:

$$.8878 + .0068 + .3819 + .0029 = 1.2794$$

Since both row and column totals are fixed, the number of

$$\text{degrees of freedom} = (r - 1)(c - 1)$$

$$= (2 - 1)(2 - 1)$$

$$= 1 \times 1 = 1$$

Entering appropriate tables of Chi square, with $\chi^2 = 1.2794$

and one degree of freedom we find

P lies between .30 and .20Calculation
checked? Yes.

Calculation
Serial No: 083

Source data sheets
not reproduced.

Comparison of the number of sheep and goats falling in:
(a) The group scoring no A hits } All judges
(b) The group scoring 1 or more A hits } uncorrected.
with the numbers to be expected on a chance basis.

Expected numbers in each category $\frac{154}{191} = .701571$ of total.
For Sheep from chance alone:

N for group sans A hits = 134 $134 \times .701571 = \underline{94.01}$

N for 1 or more group = 57 $57 \times .701571 = \underline{39.99}$

Expected numbers in each category $\frac{57}{191} = .298429$ of total.
For Goats from chance alone:

N for group sans A hits = 134 $134 \times .298429 = \underline{39.99}$

N for group 1 or more = 57 $57 \times .298429 = \underline{17.01}$

Subject Category:	No: of S's in group sans A hits	No: of S's in group scoring A hits:	Totals:
Observed	91	43	134
Sheep: Expected	94.01	39.99	
Difference	3.01	3.01	
Observed	43	14	57
Goats Expected	39.99	17.01	
Difference	3.01	5.01	
Totals	134	57	191

$$\begin{aligned} \chi^2 &= \frac{3.01^2}{94.01} + \frac{3.01^2}{39.99} + \frac{3.01^2}{39.99} + \frac{3.01^2}{17.01} \\ &= \frac{9.0601}{94.01} + \frac{9.0601}{39.99} + \frac{9.0601}{39.99} + \frac{9.0601}{17.01} \\ &= .10 + .23 + .23 + .53 = 1.09 \end{aligned}$$

Hence $\chi^2 = 1.09$ with one degree of freedom.

P lies between .50 and .20

Calculation
checked ? Yes.

(d) Conclusions from the Sheep-Goat Separations:

It will be seen from the graph presented above that the results obtained in the Rhodes experiment are in conformity with those obtained by previous investigators. The Sheep score highest, the Open-minded next highest, and the Goats lowest. Although these results are thus in the predicted direction, calculations C81 to C83 show that the differences obtained are not large enough to be regarded as significant. In C81 P is only .7 to .5. As the biggest differences were between the Sheep and Goat groups, a separate assessment was made of these categories in C82. Elimination of the Open-minded group caused the significance of the difference to improve somewhat and reach the .3 to .2 level. The slightly different method adopted in C83 yielded exactly the same probabilities as C82.

It is worth noting that the biggest contribution to the Chi-square totals, and hence to the significance of the results, comes from a failure of the Goats to score as many hits as would be expected from chance.

Schmeidler says of her findings: " The actual difference between the sheep and the goats is pitifully small. A score of 5 would of course be expected by chance; 3,909 sheep runs have a mean score of only 5.15; and the mean score of the 3,520 goat runs is 4.92. The difference between them, I have calculated roughly, is of the same order of magnitude as the house percentage in a roulette game. And I think there is an analogy here: no one can expect to run an honest roulette wheel at a profit unless he has a large reserve of funds; and probably no experimenter could expect significant differences between groups of subjects, using the sheep-goat criterion alone, unless he has a large enough reserve of patience to collect a very great many cases. " (67 page 25)

In view of this, the fact that the differences obtained in the Rhodes experiment only reach the .3 to .2 level of

significance is probably not so disappointing as would appear at first thought. To the extent to which they can be regarded as significant, therefore, they fulfil the predictions made earlier in this chapter on personality, and tend to support the findings of Schmeidler. They thus provide additional (if not very strong) evidence that the A hits scored in the experiment reflect E.S.P. ability on the part of the subjects, and are not merely the result of chance factors.

(11) Combination of Sheep-Goat Ratings with Adjustment Ratings:

(a) Introduction:

So far all the predictions made regarding the trends which the Rhodes data should show have been fulfilled. It remains now to examine the last two predictions. It will be remembered that Schmeidler found that well-adjusted subjects gave less equivocal scores than the poorly adjusted ones, and was led to suggest that if adjustment ratings were combined with the Sheep-Goat criterion the differences between the well-adjusted sheep and well-adjusted goats would be in favour of the sheep, and bigger than if only the Sheep-Goat criterion were used to separate the groups.

In the course of this chapter we have already seen that the group scoring two or more A hits are better adjusted than the group failing to score A hits, and that the sheep do better than the goats. It now remains to combine these two ratings and assess the efficiency of the combination.

(b) Statistical Methods:

It was first necessary to establish some criterion for separating the subjects into well-adjusted, averagely-adjusted, and poorly-adjusted groups. It was decided to use the Washburne sub-total score, and to divide the subjects into the worst adjusted 25 %, the best adjusted 25 %, and a middle 50% whose adjustment was average. For this purpose the

Quartile Deviation Q was calculated separately for males (C84) and females (C85), and the 25th and 75th percentiles established as the perimeters of the groups. This gave the following separations:

	Men score:	Women score:
Best adjusted 25%:	89 & less	87 & less
Middle 50%:	90 to 129	88 to 135
Worst adjusted 25%:	130 & more	136 and more.

The frequency of scoring A hits was then calculated per trial for each of the three groups. These frequencies were then tested by the now familiar Chi-square method to find whether they differed significantly from the chance expectation (C86).

Next each of these three groups were further subdivided into Sheep, Goats and Open-minded sub-groups, and the A hit frequency per trial worked out. These frequencies were in turn tested to find whether they differed significantly from chance.

(●) Calculations C 84 and C85:

Graph:2:

Calculation C 86:

Calculation
Serial No: C84

Source Data Sheets
not reproduced.

Calculation of the Quartile Deviation Q for the Washburne
Subtotal Adjustment scores for men.

Step Intervals: Frequency:

50	to	44	0
45		59	0
60		74	18
75		89	35
90		104	35
105		119	42
120		134	36
135		149	11
150		164	5
165		179	13
180		194	4
195		209	2
210		224	1
225		239	0
240		254	1
255		269	0
270		285	1
			<u>202</u>

N = 202

Step Interval = 15

Calculation of Q_1 : Quota = $\frac{N}{4} = \frac{202}{4} = 50.5$

To the beginning of step (75 to 89) gives 18 cases, leaving a deficit of 32.5 still required.

$$Q_1 = (\text{Beginning of step}) + \frac{\text{Deficit}}{\text{No. of cases in step}} \times \frac{\text{Step interval}}{1}$$

$$= 74.6 + \frac{32.5}{33} \times \frac{15}{1} = 74.6 + 14.77 = \underline{89.37}$$

Calculation of Q_3 Quota = $\frac{3N}{4} = \frac{202}{1} \times \frac{3}{4} = 151.5$

To the beginning of step (120 to 134) gives 128 cases, leaving a deficit of 23.5

$$Q_3 = 119.6 + \frac{23.5}{36} \times \frac{15}{1} = 119.6 + \frac{352.5}{36}$$

$$= 119.6 + 9.79 = \underline{129.39}$$

$$\text{Hence } Q = \frac{Q_3 \text{ minus } Q_1}{2} = \frac{129.39 - 89.37}{2} = \frac{40.02}{2}$$

$$= 20.01$$

Thus the best adjusted 25% get scores of 89 and less

The worst adjusted 25% get scores of 130 and more.

Calculation
Checked ? Yes.

Calculation
Serial No: C85.

Source Data Sheets
not reproduced.

Calculation of the Quartile Deviation Q, for the Washburne
Subtotal Adjustment scores For women.

Step Intervals: Frequency:

30	to	44	1
45		59	2
60		74	13
75		89	21
90		104	21
105		119	20
120		154	24
135		149	12
150		164	11
165		179	7
180		194	3
195		209	2
210		224	0
225		239	1
240		254	0
255		269	0
270		285	0
			<u>138</u>

$N = 138$

Step interval = 15.

Calculation of Q_1 Quota = $\frac{N}{4} = \frac{138}{4} = 34.5$ cases.

To the beginning of step (75 to 89) gives 16 cases, leaving a deficit of 18.5 cases. The step begins at 74.6

$$74.6 + \frac{18.5}{21} \times \frac{15}{1} = 74.6 + \frac{277.5}{21}$$

$$74.6 + 13.21 = \underline{87.81}$$

Calculation of Q_5 Quota = $\frac{N}{1} \times \frac{3}{4} = \frac{138}{1} \times \frac{3}{4} = 103.5$

To beginning of step (135 to 149) gives 102 cases, leaving a deficit of 1.5 cases still to get. The step begins at 134.6

$$134.6 + \frac{1.5}{12} \times \frac{15}{1} = 134.6 + \frac{22.5}{12}$$

$$134.6 + 1.875 = \underline{136.47}$$

$$\text{Hence } Q = \frac{Q_5 - Q_1}{2} = \frac{136.47 - 87.81}{2} = \frac{48.66}{2} = 24.33$$

Thus the best adjusted 25% get scores of 87 and less

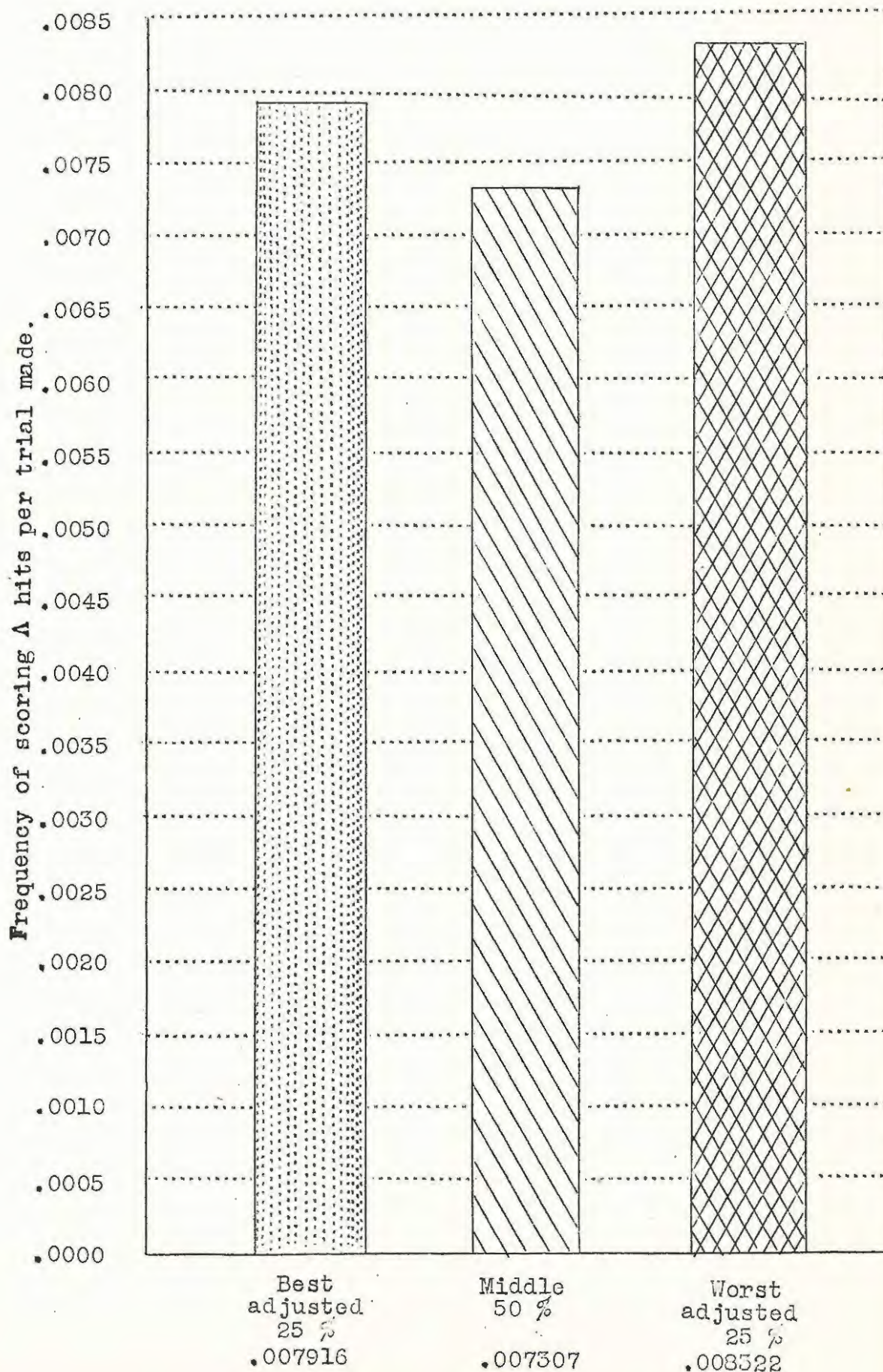
The worst adjusted get scores of 150 and above.

Calculation
checked ? Yes.

GRAPH No: 2

SHOWING THE A HIT SCORING RATE PER TRIAL (ALL JUDGES, UNCORRECTED) FOR THREE GROUPS OF SUBJECTS RATED WELL-ADJUSTED, AVERAGELY ADJUSTED AND POORLY ADJUSTED ON THE WASHBURNE SUBTOTAL SCORES.

Well-adjusted: .007,916 hits.
 Averagely-
 adjusted: .007,307 hits.
 Poorly-
 adjusted: .008,322 hits



Calculation
Serial No: C86

Source Data Sheets
not reproduced.

Examination of the numbers of subjects rated Well Adjusted, Averagely Adjusted, and Poorly Adjusted on the Washburne Subtotal Score, falling into each of the three groups:-

- | | | |
|-------------------------------------|---|-----------------------------|
| (a) Group sans A hits. | } | All judges,
uncorrected. |
| (b) Group scoring 1 A hit. | | |
| (c) Group scoring 2 or more A hits. | | |

to find whether the observed frequencies differ significantly from those to be expected on a chance basis.

Expected frequencies in each category if the numbers are distributed according to chance:

Well-adjusted: $\frac{88}{343}$ of the totals in each group. = .256,560

No A's: = 240 X .256,560 = 61.57 subjects

1 A: = 79 X .256,560 = 20.27

2 or more A's: = 24 X .256,560 = 6.16

Averagely adjusted: $\frac{164}{343}$ of the total no. of subjects in each group.
= .478,134

No A's: = 240 X .478,134 = 114.75 subjects.

1 A: = 79 37.77

2 & 2+ = 24 11.47

Poorly adjusted: $\frac{91}{343}$ of the totals in each group. = .265,306

No A's: 240 X .265,306 = 63.67 subjects.

1 A: 79 20.96

2 & 2+ 24 6.37

Subject category:	No A Hits	1 A Hit	2 or more	Totals	
Well-adjusted	Obs.	62	19	7	88
	Exp.	61.57	20.27	6.16	
	Diff.	.45	1.27	.84	
Averagely adjusted	Obs.	118	34	12	164
	Exp.	114.75	37.77	11.47	
	Diff.	3.25	3.77	.53	
Poorly adjusted	Obs.	60	26	5	91
	Exp.	63.67	20.96	6.37	
	Diff.	3.67	5.04	1.37	
Totals:	240	79	24	343	

Calculation

Serial No: C86. continued. page 2.

The next step is to find the Chi-square value for each of the nine cells in the table as follows:-

$$\frac{.43^2}{61.57} = .00$$

$$\frac{1.27^2}{20.27} = .08$$

$$\frac{.84^2}{6.16} = .11$$

$$\frac{3.25^2}{114.75} = .09$$

$$\frac{3.77^2}{37.77} = .38$$

$$\frac{.53^2}{11.47} = .02$$

$$\frac{3.67^2}{65.67} = .21$$

$$\frac{5.04^2}{20.96} = 1.21$$

$$\frac{1.37^2}{6.37} = .29$$

$$\underline{2.39} \quad \text{Total } \chi^2$$

Since both row and column totals are fixed, the degrees of freedom are given by the formula $(r - 1)(c - 1)$

$$= (3 - 1)(3 - 1) = 2 \times 2 = \underline{4 \text{ d.f.}}$$

Entering appropriate tables of Chi-square with $\chi^2 = 2.39$ and 4 degrees of freedom gives:

P lies between .70 and .50

Calculation
checked ? Yes.

(d) Comments on the preliminary adjustment data:

It will be remembered that the mean-difference analysis brought to light a tendency for the group scoring 2 or more A hits to be better adjusted (in terms of Washburne Sub-total scores) than other groups, although this difference was not large enough to be statistically significant. On the basis of the work of previous investigators it had been predicted that the data would show such a trend.

The results of the latest breakdown of data do not conform to expectations. As will be seen from the graph the worst adjusted subjects show the highest scoring rate. Calculation C 86 shows however that the differences are not significant. This reversal in trend is probably due to the inclusion of subjects scoring 1 A hit into the hit scoring group, and the greater dilution from chance hits thus introduced.

Consideration was given to using the group scoring 2 or more A hits alone, but a preliminary breakdown of the data gave the following figures:

	<u>Well adjusted</u>	<u>Average</u>	<u>Poorly adjusted:</u>
No. of subjects observed in group:	7	12	5
Expected numbers:	6	12	6

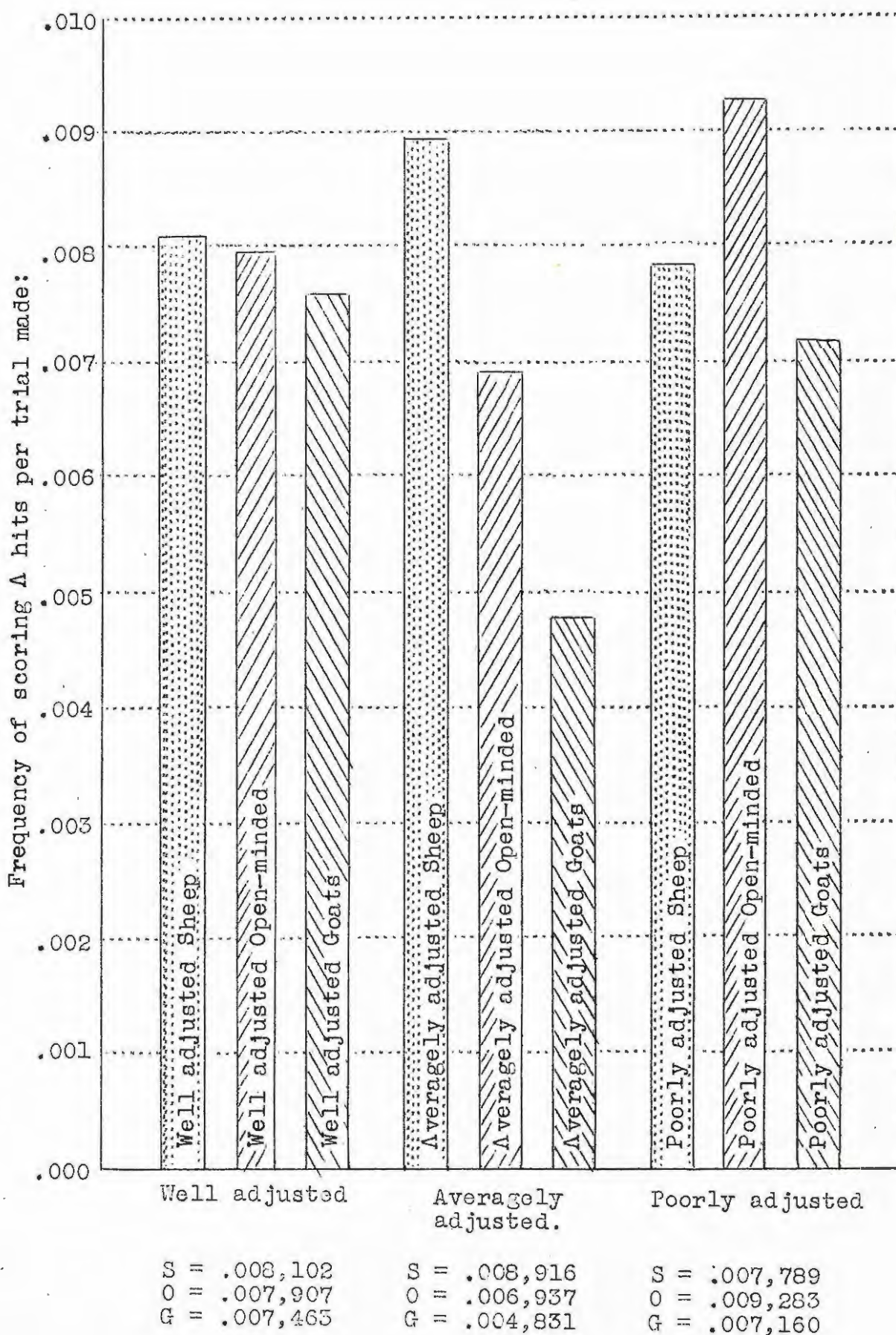
From inspection it seems clear that no significant difference exists.

Although the preliminary examination of the Washburne adjustment rating separations does not look at all promising, there is a faint possibility that differential scoring on the part of Sheep and Goats is obscuring the real trends. Consequently the next step will be to break down the data further into Sheep, Open-minded and Goat categories, and calculate the frequencies of A hit per trial for each of these groups.

(e) Graph: 3:

Calculation C 87:

GRAPH No: 3

COMBINATION OF SHEEP-GOAT RATINGS WITH
WASHBURNE SUBTOTAL ADJUSTMENT RATINGS.SHOWING NUMBER OF A HITS (ALL JUDGES:
UNCORRECTED) SCORED PER TRIAL FOR EACH
SUBJECT CATEGORY.

Calculation
Serial No: C'87.

Source Data Sheets
not reproduced.

Calculation to find whether the 9 subject categories obtained by classifying subjects as "Sheep", "Goats" & "Open-minded", and "Well-adjusted", "Averagely-adjusted" and "Poorly-adjusted" show frequencies of scoring A hits per trial which differ significantly from those to be expected from chance.

Expected frequencies:

Well-adjusted:

Sheep:	$\frac{1,728}{16,679} = .103,603$	X 129 = 13.36 hits
Open-minded:	$\frac{1,897}{16,679} = .113,736$	14.67
Goats:	$\frac{670}{16,679} = .040,170$	5.18

Averagely-adjusted:

Sheep:	$\frac{2,804}{16,679} = .168,115$	21.69
Open-minded:	$\frac{3,892}{16,679} = .233,347$	30.10
Goats:	$\frac{1,242}{16,679} = .074,465$	9.61

Poorly-adjusted:

Sheep	$\frac{1,669}{16,679} = .100,066$	12.91
Open-minded:	$\frac{1,939}{16,679} = .116,254$	15.00
Goats:	$\frac{838}{16,679} = .050,243$	6.48

continued:

Calculation
Serial No: 007, continued, page 2.

Subject category		Hits	Misses	Totals:
Well adjusted Sheep	Obs.	14	1,714	1,728
	Exp.	13.36	1,714.64	
	Diff.	.64	.64	
Well- adjusted Open-minded:	Obs.	15	1,882	1,897
	Exp.	14.67	1,882.33	
	Diff.	.33	.33	
Well- adjusted Goats:	Obs.	5.	665	670
	Exp.	5.18	664.82	
	Diff.	.18	.18	
Averagely- adjusted Sheep:	Obs.	25	2,779	2,804
	Exp.	21.69	2,782.31	
	Diff.	3.31	3.31	
Averagely- adjusted Open-minded:	Obs.	27.	3,865	3,892
	Exp.	30.10	3,861.90	
	Diff.	3.10	3.10	
Averagely- adjusted Goats:	Obs.	6	1,236	1,242
	Exp.	9.61	1,232.39	
	Diff.	3.61	3.61	
Poorly- adjusted Sheep:	Obs.	13.	1,656	1,669
	Exp.	12.91	1,656.09	
	Diff.	.09	.09	
Poorly- adjusted Open-minded:	Obs.	18	1,921	1,939
	Exp.	15.00	1,924.00	
	Diff.	3.00	3.00	
Poorly- adjusted Goats:	Obs.	6.	832.	838
	Exp.	6.48	831.52	
	Diff.	.48	.48	
Totals:		129	16,550	16,679

Calculation
Serial No: C87, continued, page 3.

Chi-square totals for each cell are given by the formula:

$$\frac{\text{Difference}^2}{\text{Expected frequency.}}$$

	<u>Hits:</u>		<u>Misses:</u>
WAS	$\frac{.64^2}{13.36} = \frac{.4096}{13.36} = .031$		$\frac{.64^2}{1,714.64} = \frac{.4096}{1,714.64} = .0002$
WAO	$\frac{.33^2}{14.67} = \frac{.1089}{14.67} = .007$		$\frac{.33^2}{1,882.33} = \frac{.1089}{1,882.33} = .0001$
WAG	$\frac{.18^2}{5.18} = \frac{.0324}{5.18} = .006$		$\frac{.18^2}{664.82} = \frac{.0324}{664.82} = .0000$
AAS	$\frac{3.31^2}{21.69} = \frac{10.9561}{21.69} = .505$		$\frac{3.31^2}{2,782.31} = \frac{10.9561}{2,782.31} = .0039$
AAO	$\frac{3.10^2}{30.10} = \frac{9.61}{30.10} = .319$		$\frac{3.10^2}{3,861.90} = \frac{9.61}{3,861.90} = .0025$
AAG	$\frac{3.61^2}{9.61} = \frac{13.0321}{9.61} = 1.356$		$\frac{3.61^2}{1,232.39} = \frac{13.0321}{1,232.39} = .0106$
PAS	$\frac{.09^2}{12.91} = \frac{.0081}{12.91} = .001$		$\frac{.09^2}{1,656.09} = \frac{.0081}{1,656.09} = .0000$
PAO	$\frac{3.00^2}{15.00} = \frac{9.00}{15.00} = .600$		$\frac{3.00^2}{1,924.00} = \frac{9.00}{1,924.00} = .0047$
PAG	$\frac{.48^2}{6.48} = \frac{.2304}{6.48} = .035$		$\frac{.48^2}{831.52} = \frac{.2304}{831.52} = .0003$

Total for whole table: 2.8823

Degrees of freedom for the whole table = $(r - 1)(c - 1)$
 $= (9 - 1)(2 - 1) = 8 \times 1 = \underline{\text{Eight degrees of freedom.}}$

Entering appropriate tables with these values

P lies between .95 and .90

Calculation
checked ? Yes.

(f) Comments on the combined Sheep-Goat &
Adjustment Separations:

Calculation C87 shows that the A hit frequencies shown by the various groups do not depart significantly from those to be expected from chance. P for the whole table lies between .95 & .9, which is clearly not significant.

However the data shows certain trends that need comment. As predicted by Schmeidler the well-adjusted sheep do better than the well-adjusted goats. Contary to the last prediction, however, breaking down the sheep-goat categories has produced a smaller difference than that obtained for all the sheep and all the goats. (.000,639 as against .002,225 per trial.)

There is a general trend for the sheep to do best, the open-minded to do next best, and the goats worst of all for all the adjustment categories, the one exception being the poorly-adjusted open-minded class. It will be remembered that Schmeidler warned that less clear-cut distinctions could be expected from the poorly-adjusted groups.

Thus to the extent to which the results can be regarded as significant, they fulfil the fourth prediction, but not the fifth one.

(12) General Conclusions from the Personality Test Material:

Before undertaking the analysis of the personality test material five predictions were made as to the trends the data should show if produced by genuine E.S.P. ability on the part of the subjects, and not merely by chance factors. In this review of these predictions and the findings bearing on them, we shall consider only the differences between the group failing to score A hits, and the group scoring two or more A hits.

Since we are predicting not only that certain differences will be found, but also the direction of the differences, it will be in order to use a one-tailed test of significance, and the P values given below are all for such a one-tailed test. (If any reader prefers the two-tailed test, the P values should be doubled).

Prediction 1: If the subjects are divided into a high-scoring group, and a low-scoring group, the high scoring subjects will be found to be more extraverted in terms of the Bernreuter B3I scores than the low scoring group.

Findings: Assessment by the mean-difference method showed the high scorers to be significantly more extraverted, P (one-tail) lying between .0005 and .0003. The median test confirmed this trend, but at a lower level of significance, P (one-tail) dropping to between .05 and .025.

Prediction 2: The high scoring subjects will be found to be better adjusted than the low scoring group.

Findings: Three possible measures of adjustment were assessed: the Washburne sub-total, the Bernreuter F1-C scale (since it correlates +.95 with the B1-N Neurotic Tendency scale), and the B3-I scale (since it too correlates +.90 with B1-N).

As far as the Washburne Subtotal was concerned, the mean difference assessment showed the predicted trend, but with P (one-tail) only reaching the .25 level of significance. When later a different assessment was made, in terms of Q, the

trend was reversed.

The measure of adjustment given by the Bernreuter F1-C scores showed the predicted trend in both mean difference and median test assessments, P (one tail) reaching .0244 and .025 to .01 respectively.

Using the Bernreuter B3-I scores, both mean difference assessments and median test assessments showed the predicted trend, P (one-tail) reaching .0005 to .0003 in the mean difference assessment, and .05 to .025 in the median test.

Prediction 3: The subjects who can be classified as "sheep" in the Rhodes experiment should score more hits than the "goats", though this difference can be expected to be relatively small.

Findings: The sheep scored .008,301 hits per trial as against the .006,076 h.p.t. scored by the goats. This is a difference of .002,225 h.p.t. A test of significance gave a P (one tail) lying between .15 and .10. Thus the trend is shown, but fails to reach the .1 level of significance.

Prediction 4: If "sheep" and "goats" are further classified as "well-adjusted" and "poorly-adjusted", the "well-adjusted sheep" will score more hits than the "well-adjusted goats".

Findings: Here again the predicted trend was shown, but failed to reach the .1 level of significance. The "well-adjusted sheep" scored .008,102 h.p.t. as against the "well-adjusted goats" .007,463 h.p.t., giving a difference in the predicted direction of .000,639 h.p.t.

Prediction 5: The differences in scoring rates between the "well adjusted sheep" and the "well-adjusted goats" will be greater than that between the "sheep" and "goat" groups as a whole.

Findings: The predicted trend was not shown in the Rhodes results, the difference between the "sheep" and "goats" as a whole being .002,225 as against .000,639 h.p.t. for the groups of "well-adjusted sheep" and "well-adjusted goats"

We have thus found that in the case of the A hits, the Rhodes data shows four out of the five predicted trends. In the case of those trends reaching statistical significance, a number of alternative hypotheses were considered to the E.S.P. one, but were found inadequate to account for the results. Since none of the counter-hypotheses discussed in the First Main Analysis are capable of accounting for them, unless the reader can think of additional hypotheses overlooked by the writer, these significant personality relationships which have been found must be taken as establishing that E.S.P. on the part of the subjects has been a factor in producing the A hits.

The discovery of these significant personality relationships in the Rhodes data thus removes any ambiguity in the interpretation of the results of the First Main Analysis.

It is perhaps also worth observing that although nothing in the Rhodes results contradicts the findings of Schmeidler regarding the "sheep-goat" ratings, and their combination with "adjustment" ratings, it would seem from the Rhodes findings that these trends are very slight, and are easily obscured by chance factors. It may be that as Schmeidler used Rorschach ratings for adjustment separations, and worked with E.S.P. cards rather than drawings, the Rhodes data is not strictly comparable to hers.

We have now come to the end of our consideration of the personality test material, and can turn to a consideration of the Psychic Questionnaire. This material will form the concluding section of the Third Main Analysis.

(13) Comparison of certain answers to the Psychic Questionnaire returned by a group scoring 2 or more A hits, with answers returned by a group scoring 0 A hits, to discover if any differences revealed are statistically significant:

(a) Introduction:

It would obviously be of great value to be able to predict from replies to a questionnaire which subjects were likely to make good scores on E.S.P. tests, and which bad scores. To this end it is proposed now to examine some of the replies made to questions in the Psychic Questionnaire by two groups of subjects, the one group consisting of the most successful subjects in the Rhodes experiment, namely those scoring 2 or more A hits, and the least successful group of subjects failing to score any A hits at all. It is intended to try and discover if any question is answered one way by the first group, and a different way by the second. If any such questions are discovered, the next step will be to assess the differences to find if they are large enough to be statistically significant.

(b) Statistical method:

The first step is simple compilation. As example, let us consider question 1 in the Imagery section of the questionnaire, reading as follows:

"How frequently do you dream ? (a) Very seldom
(b) Once a month
(c) Once a week
(d) Every few nights
(e) Nightly.

First we consider the group scoring two or more A hits, 24 in number. An analysis of their replies shows that 4 answered (a), 2 answered (b), 1 answered (c), 14 answered (d) and 3 answered (e). (These figures are set out in Table XI below.). In the same way frequencies are determined for all the subjects in the group failing to score A hits who answered question 1. (Not all subjects answered all questions, so that the totals vary somewhat from question to question). In this case, 248 subjects answered. The distribution of their answers was found to be as follows:

67 answered (a), 18 answered (b), 36 answered (c), 78 answered (d) and 49 answered (e).

To assess the significance of the differences between these two groups it is proposed to use the χ^2 test for two independent samples as described by Siegel. (197, p. 104). This requires that the data be cast into a 2 X 2 table. This test requires that no expected frequency in any cell fall below 5, so that in questions with more than two possible replies it has been thought advisable to combine adjacent categories. In the example on dream frequencies we are considering, this results in the following 2 X 2 contingency table:

	E.S.P. Group.	Non-E.S.P. Group:	Totals:
Dream frequencies:			
Once a week or less (a,b,c)	7	121	128
Oftener than once a week (d,e,)	17	127	144
Totals:	24	248	272

The null hypothesis we are going to test is that there is no difference in dream frequencies between the E.S.P. group, and the Non-E.S.P. group.

The χ^2 test for two independent samples is chosen because the two groups, E.S.P. and Non-E.S.P. subjects, are independent, and the "scores" under study are frequencies in discrete categories: frequent and infrequent dreamers.

The same formula used in the median tests is again used here, namely:

$$\chi^2 = \frac{N(|AD - BC| - \frac{N}{2})^2}{(A+B)(C+D)(A+C)(B+D)}$$

χ^2 computed from this formula has a sampling distribution which is approximated by the χ^2 distribution with $df = 1$.

Since we have slight grounds if any for predicting the direction any differences will take, a two-tailed test of significance is appropriate.

The application of this formula is in no way different from its application in the median tests already carried out, so this will not be detailed again here.

From Tables XI and XII it will be seen that no probability figures are given for many questions. This is because the differences obtained are either so small that their insignificance can be assumed, or else where the differences are promising, the theoretically expected frequencies in one cell have dropped below 5, rendering the χ^2 test misleading. (An example of this is "Sleepwalking as an adult").

The difficulty in this second case is that the results fall between two statistical stools. Because the E.S.P. group is so small, only 24, the expected frequencies for this group often fall below 5, rendering it invalid to apply the χ^2 test. At the same time, because the group failing to score A hits is so big, making the total N so large, it becomes prohibitively laborious to apply the Fisher Exact Probabilities Test. (107, p. 96). This can be shown by taking the data on sleepwalking as an adult as an example.

		E.S.P. group.	Non-E.S.P. group.	Totals:
Sleepwalking as an adult:	Yes	A 7	B 27	34
	No	C 14	D 214	228
Totals:		21	241	262

Here the expected frequency in cell A falls below 5. To calculate the significance by the Fisher Exact Probabilities Test would involve the application of the following formula:

$$p = \frac{(A + B)! (C + D)! (A + C)! (B + D)!}{N! A! B! C! D!}$$

This formula is applied to the data in the above table, and then to each of the more extreme values of A possible, namely A=6, B=28; A=5, B=29; A=4, B=30; A=3, B=31; A=2, B=32; A=1, B=33; and A=0, B=34. In view of the labour involved in such a computation, the significance for any question where the expected frequency in any cell falls below 5 has not been calculated, even though the differences appear promising.

Table XI below shows the results when subjects returning less than the full 50 drawings are included in the groups, Table XII shows the results when all subjects returning less than the full 50 drawings are purged from the groups. Table XIII compares the results of the two previous tables.

(c) Calculations C 88 to C 103:

Tables XI, XII and XIII:

Calculation
Serial No.: C 88.

Source data can be obtained
from Table XI.

DREAM FREQUENCIES:

Data:	E.S.P. Group		Non- E.S.P. Group	Totals:
Once a week or less	A	7	B 121	128
Oftener than once a week:	C	17	D 127	144
Totals:		24	248	272 = N

Formula: $\chi^2 = \frac{N(|AD - BC| - \frac{1}{2}N)^2}{(A+B)(C+D)(A+C)(B+D)}$ (The $-\frac{1}{2}N$ in the top line is a correction for continuity.)

Smallest expected frequency should be greater than 5. It is:

$$E_A = \frac{24 \times 128}{272} = > 5$$

$$(A+B)(C+D)(A+C)(B+D) = 128 \times 144 \times 24 \times 248 = 109,707,264$$

$$|AD - BC| = 889 - 2,057 = 1,168$$

$$1,168^2 = 1,364,224$$

$$1,168^2 \times N = 371,068,928$$

$$-\frac{1}{2}N = 136 \quad 1,168 - 136 = 1,032$$

$$1,032^2 = 1,065,024$$

$$1,032^2 \times N = 289,686,528$$

Uncorrected χ^2 (with 1 d.f.):

$$\frac{371,068,928}{109,707,264}$$

3.38

P lies between: .1 and .05

Corrected χ^2 (with 1 d.f.):

$$\frac{289,686,528}{109,707,264}$$

2.64

P lies between: .2 and .1

Calculation
checked ? Yes.

TABLE XI:

SUMMARY OF DATA FROM PSYCHIC QUESTIONNAIRE: PAGE I.
 (INCLUDING SUBJECTS FAILING TO RETURN THE FULL 50
 DRAWINGS).

(Calculations C88 to C96)

Question:	Ans. Sym.	Group scoring 2 or more A type hits:		Group failing to score A type hits:		P. (2-tail)
		Score:	%	Score	%	
Dream frequencies:	a	4	16.67	67	27.016	.1-.05
	b	2	8.33	18	7.258	
	c	1	4.17	36	14.516	
	d	14	58.33	78	31.452	
	e	3	12.50	49	15.758	
	T	24		248		
Hypnagogic Imagery experienced	Y	18	78.26	188	77.05	
	N	5	21.74	56	22.95	
	T	23		244		
Aware of friends' moods	Y	22	95.65	226	91.87	
	N	1	4.35	20	8.13	
	T	23		246		
Aware of strangers' moods:	Y	12	52.17	143	58.85	
	N	11	47.83	100	41.15	
	T	23		243		
Aware of conversation undercurrents	Y	21	91.30	224	90.32	
	N	2	8.70	24	9.68	
	T	23		248		
Aware of atmosphere	Y	22	95.65	203	82.52	Too small
	N	1	4.35	43	17.48	
	T	23		246		
Experience of "slipping away":	Y	18	81.82	172	69.92	.3-.2 .5-.3
	N	4	18.18	74	30.08	
	T	22		246		
Sleepwalking as a child:	F	1	4.35	10	4.12	.5-.3 .7-.5
	O	7	30.43	53	21.81	
	N	15	65.22	180	74.07	
	T	23		243		
Sleepwalking as an adult:	F	0	0	2	.83	Too small
	O	7	33.33	25	10.37	
	N	14	66.67	214	88.80	
	T	21		241		
Sleeptalking as a child:	F	6	27.27	46	19.25	.05-.02 .1-.05
	O	15	68.18	134	56.07	
	N	1	4.55	59	24.68	
	T	22		239		
Sleeptalking as an adult:	F	2	8.70	27	11.20	.1-.05 .2-.1
	O	18	78.26	139	57.68	
	N	3	13.04	75	31.12	
	T	23		241		

Continued:

TABLE XI: PAGE 2:

Question:	Ans. Sym.	Group scoring 2 or more A type hits:	Group failing to score A type hits:	P. (2-tail)		
Experience of Deperson.	Y	7	30.43	114	46.53	.2 -.1
	N	16	69.57	131	53.47	.3 -.2
	T	23		245		
Religious denomination:	Ang	12	52.17	129	52.44	
	Meth	1	4.35	23	9.35	
	Cath	1	4.35	13	5.28	
	Pres	4	17.39	23	9.35	
	Bapt	0	0	4	1.63	
	C Sc	1	4.35	4	1.63	
	Luth	0	0	2	.81	
	DRC	1	4.35	15	6.10	
	Spir	0	0	1	.41	
	Jew	0	0	21	8.54	
	Nil	3	13.04	11	4.47	
T	23		246			
Freq. of church attendance	W	6	26.09	99	42.13	.2 -.1
	M	6	26.09	30	12.77	.3 -.2
	Y	8	34.78	84	35.74	
	N	3	13.04	22	9.36	
	T	23		235		
Mystical exper.	Y	1	4.35	21	8.50	Too small
	?	4	17.39	11	4.45	
	N	18	78.26	215	87.04	
	T	23		247		
Personal psychical exper.	Y	14	60.87	147	60.49	
	N	9	39.13	96	39.51	
	T	23		243		
Phenomena accepted "B" or "E"	Y	22	95.65	204	83.61	Too small
	N	1	4.35	40	16.39	
	T	23		244		
Déjà Vu experience	Y	18	78.26	163	67.35	.3 -.2
	N	5	21.74	79	32.64	.5 -.3
	T	23		242		
Relations cited as psychic:	Y	3	13.04	24	10.26	Too small
	N	20	86.96	210	89.74	
	T	23		234		
Self cited as psychic:	Y	2	8.70	10	4.31	
	N	21	91.30	222	95.69	
	T	23		232		
Hallucinations exper.	M	6	26.09	41	17.30	.2 -.1
	O	2	8.70	9	3.80	.3 -.2
	N	15	65.21	187	78.90	
	T	23		237		

P values have been obtained from Siegel, page 249, and are for a two-tailed test of χ^2 with 1 d.f. The first P value given is not corrected for continuity, the second is with the correction for continuity applied.

Transcription checked ?
Yes.

TABLE XII:

SUMMARY OF DATA FROM PSYCHIC QUESTIONNAIRE: PAGE 1.
(EXCLUDING SUBJECTS FAILING TO RETURN THE FULL 50
DRAWINGS).

(Calculations C97 to C103).

Question:	Ans. Sym.	Group scoring 2 or more A type hits:		Group scoring 0 A type hits:		P (2-tail)		
		Score:	%	Score:	%			
Dream frequencies	a	4	7	19	93	26	.3 -.2	
	b	2		10		7		.5 -.3
	c	1		5		14		
	d	11	14	52	104	32		
	e	3		14		20		
	T	21		197				
Hypnagogic Imagery experienced	Y	16	80	151	77			
	N	4	20	44	23			
	T	20		195				
Aware of friends' moods:	Y	19	95	180	92			
	N	1	5	16	8			
	T	20		196				
Aware of strangers' moods:	Y	10	50	113	59			
	N	10	50	79	41			
	T	20		192				
Aware of conversation under- currents:	Y	18	90	180	91			
	N	2	10	17	9			
	T	20		197				
Aware of atmosphere	Y	19	95	166	85	Too small		
	N	1	5	29	15			
	T	20		195				
Experience of "slipping away":	Y	16	84	137	70	.2 -.1 .3 -.2		
	N	3	16	59	30			
	T	19		196				
Sleepwalking as a child:	F	1	7	5	49	4	.5 -.3	
	O	6		30		21		.7 -.5
	N	13	65	144	75			
	T	20		193				
Sleepwalking as an adult:	F	0	6	0	19	.5	Too small.	
	O	6		32		18		9.5
	N	13	68	172	90			
	T	19		191				
Sleeptalking as a child:	F	6	13	32	155	18	Too small.	
	O	12		63		110		58
	N	1		5		45		24
	T	19		190				
Sleeptalking as an adult:	F	2	17	10	132	10	.2 -.1 .2 -.1	
	O	15		75		113		58
	N	3		15		61		32
	T	20		193				

Continued:

TABLE XII: PAGE 2:

Question:	Ans. Sym.	Group scoring 2 or more A type hits:		Group failing to score A type hits:		P (χ^2 -tail)
		Score:	%	Score:	%	
Experience of depersonalization:	Y	7	35	97	50	.3 - .2 .5 - .3
	N	13	65	98	50	
	T	20		195		
Religious denomination:	Ang	9	45	109	55	
	Met	1	5	16	8	
	Cat	1	5	10	5	
	Pre	4	20	16	8	
	Bap	0	0	3	1.5	
	CSc	1	5	3	1.5	
	Lu	0	0	2	1	
	DRC	1	5	11	6	
	Spi	0	0	1	.5	
	Jew	0	0	16	8	
	Nil	3	15	10	5	
T	20		197			
Church attendance frequency	W	5	25	86	45	.1 - .05 .2 - .1
	M	6	30	21	11	
	Y	7	35	63	33	
	N	2	10	20	11	
	T	20		190		
Mystical experience	Y	0	0	14	7	
	?	4	20	11	5.5	
	N	16	80	172	87	
	T	20		197		
Personal psychic experience	Y	12	60	120	62	
	N	8	40	74	38	
	T	20		194		
Phenomena accepted "B" or "E"	Y	19	95	158	81	Too small
	N	1	5	37	19	
	T	20		195		
Déjà Vu experience	Y	17	85	131	68	.2 - .1 .2 - .1
	N	3	15	62	32	
	T	20		193		
Relations cited as psychic:	Y	2	10	19	10	
	N	18	90	167	90	
	T	20		186		
Self cited as psychic	Y	1	5	10	5	
	N	19	95	174	95	
	T	20		184		
Hallucinations experience	M	5	25	32	17	Too small
	O	2	10	8	4	
	N	13	65	149	79	
	T	20		189		

Probability values have been obtained from Siegel, p. 249, and are for a two-tailed test of significance for χ^2 with 1 df. The first P value given is for χ^2 uncorrected for continuity. The second value is the probability when the correction for continuity has been applied.

Transcription checked ?
Yes.

TABLE XIII:

COMPARISON OF P VALUES FROM TABLES XI AND XIII TO
REVEAL THE EFFECT OF PURGING THE GROUPS OF ALL THE
SUBJECTS FAILING TO RETURN THE FULL FIFTY DRAWINGS

Question on:	Table XI P values All subjects group	Table XII P values: Purged group:
Dream frequencies:	.1 to .05 .2 .1	.3 to .2 .5 .3
Slipping away:	.3 .2 .5 .3	.2 .1 .3 .2
Sleepwalking as child:	.5 .3 .7 .5	.5 .3 .7 .5
Sleeptalking as child:	.05 .02 .1 .05	Too small.
Sleeptalking as adult:	.1 .05 .2 .1	.2 .1 .2 .1
Experience of depersonalization:	.2 .1 .3 .2	.3 .2 .5 .3
Church attendance:	.2 .1 .3 .2	.1 .05 .2 .1
Experience of Déjà Vu	.3 .2 .5 .3	.2 .1 .2 .1
Experience of hallucinations:	.2 .1 .3 .2	Too small.

Transcription checked ?
Yes.

(d) Conclusions drawn from Tables XI, XII and XIII:

Reference to these tables will show that a pair of P values is given for each assessment. The first of these has not been corrected for continuity, the second has, as Siegel recommends. We will be referring to the corrected P values throughout this discussion, all being for a two-tailed test of significance.

Table XIII shows that when all subjects are included in the groups, the most promising question is that sampling sleep-talking experiences as a child. Of the E.S.P. group, only 4.55% had not talked in their sleep in childhood, as against 24.68% in the Non-E.S.P. group. The P for this difference lay between .1 and .05.

Experience of sleep-talking as an adult did not show quite such marked differences, 13.04% of the E.S.P. group having never experienced it as against 31.12% in the Non-E.S.P. group. P here lay between .2 and .1.

70.83% of the E.S.P. group dream frequently as against only 47.21 of the Non-E.S.P. group who do so. Again P = .2 to .1.

When the groups were purged of subjects failing to return the full 50 drawings, three relationships significant at the .2 to .1 level were found, namely: (a) experience of sleep-talking as an adult, only 15% of the E.S.P. group having failed to experience it as against 32% in the Non-E.S.P. group; (b) Frequent church attendance, only 25% of the E.S.P. group attending frequently as against 45% in the Non-E.S.P. group; and (c) Experience of Déjà vu, 85% of the E.S.P. group having experienced this phenomena as against only 68% in the Non-E.S.P. group.

Of the promising differences for which no assessment of significance could be obtained, the following should be mentioned: (a) The E.S.P. group were more aware of atmosphere, (b) as adults they sleep-walked more frequently (c) they accepted the reality of psychic phenomena more frequently, and (d) they more frequently experienced hallucinations than did the Non-E.S.P group.

It should however be emphasised that none of the differences assessed reached the .05 level of significance. Although the data may suggest the relationships outlined above, it does not establish them. In spite of this, the analysis may nevertheless be of assistance to future investigators by drawing their attention to the questions which are promising, and encouraging them to undertake a further investigation of these with larger groups of subjects.

This completes our examination of the Personality Material, and we are now free to turn our attention to the question of Linkage.

FOURTH MAIN ANALYSIS:LINKAGE.

(1) <u>Introduction:</u>	270
(a) The nature of the problem:	270
(b) Linkage concepts encountered in anthropological studies of magic:	271
(c) Linkage concepts encountered in psychical research:	272
(d) Review of the experimental procedure:	282
(2) <u>Graphical presentation of the data:</u>	287
(a) Adjustments to Raw Scores:	287
(b) Graphs:	288
(c) Trends in scoring rate revealed by an inspection of the graphs:	292
(3) <u>General Discussion on the choice of suitable statistical techniques for the linkage assessments</u>	294
(4) <u>Assessment of the significance of the Linkage Score fluctuations by averaging the hits from all weeks pooled to obtain theoretically expected frequencies for a χ^2 test.</u>	298
(a) Statistical Method:	298
(b) Calculations C 104 to C 131:	301
Table XIV:	302
(c) Conclusions drawn from Table XIV:	303
(5) <u>Assessment of the significance of the Linkage Score Fluctuations by eliminating week B, and averaging the hits in weeks A, C, D and E to obtain theoretically expected frequencies for a χ^2 test:</u>	303
(a) Statistical Method.	303
(b) Calculations C 132 to C 171:	304
Tables XV and XVI	305
(c) Conclusions drawn from Tables XV and XVI	307

FOURTH MAIN ANALYSIS:LINKAGE.

(6)	<u>Significance of the differences in scoring rate observed between week A and each other week in the experiment:</u>	308
	(a) Statistical Method, including a discussion of when Yates' Correction for Continuity should be applied.	308
	(b) Calculations C 172 to C 211:	311
	Tables XVII and XVIII	312
	(c) Conclusions drawn from Tables XVII and XVIII:	314
	i) Week B: Colour squares	315
	ii) Week C: Handkerchief	315
	iii) Week D: Photograph & personality material	316
	iv) Week E: Photograph & personality material with agent using colour squares.	316
(7)	<u>Comparison of the scoring rates for E.G.S. with those of the C.G.S. on the E.S.O. only, to find whether they show significant differences:</u>	319
	(a) Introduction:	319
	(b) Statistical Method:	319
	(c) Calculations C 212 to C 218:	320
	Table XIX:	321
	(d) Conclusions drawn from Table XIX:	322
(8)	<u>Summary of the conclusions drawn from the Fourth Main Analysis of the Linkage data:</u>	323

FOURTH MAIN ANALYSIS: LINKAGE.(1) Introduction:(a) The nature of the problem:

One of the most interesting problems confronting parapsychologists today is the nature of the relationship established between the agent and the subject in a psi experiment. In what sense does the subject establish contact with the agent, and vice versa, and how is the subject enabled to select the particular agent, and the particular material in his mind that is relevant to the experiment being performed? This problem has been called the linkage problem in this thesis.

It is of course well known that certain conditions tend to favour telepathic exchange, whereas other conditions inhibit it. But little is known about the way in which these factors exert their influence.

It is as well to bear in mind, throughout the discussion that follows, that the problem may be an unreal one. It is true that the brains of the agent and the subject may be separated in space, leading to the feeling that the gap has to be bridged in some way before any psi action can take place. But Broad (10) has shown that brain to brain communication is only one of a number of possible hypotheses. The relationship might be mind to mind. The assumption that minds have spacial qualities and are located in a particular place may not be valid. Some psychologists, for example, have employed such concepts as the group mind and the racial unconscious, that imply a mental organisation that transcends the minds of individuals. Broad himself, at the end of an analysis that clarifies very considerably the philosophical aspects of normal and paranormal cognition, is led to introduce for serious consideration "the possibility that a person's experiences initiate more or less permanent modifications of structure or process in something which is neither his mind

nor his brain. There is no reason to suppose that this Substratum would be anything to which possessive adjectives such as "mine" and "yours" and "his" could properly be applied, as they can be to minds and to animated bodies"

Since such postulates make use of a shared substratum the question of linkage between agent and subject is clearly meaningless if they are correct.

Thouless and Wiesner (92) have also contributed importantly to clarify the issues involved, but it is not the writer's intention to consider such philosophical issues further here. Instead it is proposed to review some of the techniques which are held to bring about linkage between subject and agent, and then show how some of these techniques have been brought to experimental test in this experiment.

(b) Linkage concepts encountered in anthropological studies of magic:

A study of the anthropological material of many different cultures reveals widespread beliefs that certain procedures help to establish links between people, which make possible psi action between them. Thus Margaret Mead (43) reports that the Arapesh believe that to harm a man, a sorcerer must be in possession of "dirt" emanating from him. This "dirt" is of two kinds:

External dirt: such as parts of food, half-smoked cigarettes, butts of sugar cane.

Emanations from the body: such as perspiration scabs, saliva, semen and vaginal secretions (but excluding excreta).

The Arapesh go to great trouble to dispose of such dirt in such a way that it will not fall into the hands of sorcerers. The possession of such dirt enables the sorcerer to establish contact with his victim, and influence him.

Similar beliefs could be quoted from many other cultures. In some cases it is held that if such material is harmed, the harm will be transmitted to the man from whom the dirt came. Here the concept underlying such procedures is best summed up by Frazer's law of contagious magic: "Things that have once been in contact with each other, continue to act on each other at a distance after the physical contact has been severed." (21)

A second set of procedures centre round the creation of a symbol to stand for the person the magician wishes to contact or influence. In some cases the symbol is concrete; a doll or an animal, and is made as much like the original it represents as possible, by being dressed similarly, or by means of a mock baptismal service it is endowed with the name of the person it is desired to harm. In other cases the symbol is purely verbal --- the person whom it is desired to contact being carefully described and named. (41).

A third way contact may be brought about is believed to be through the intermediation of spiritual beings, such as the familiars of Western European witches (41, 46), and their South African counterparts, the Pondo witches (38) and Zulu sorcerers (40).

The first two of these procedures can be termed magical, since they are thought to depend on a knowledge of laws governing supernatural forces. The third could be classified as religious, since it is a projection of attitudes towards people, and not attitudes towards things, as in magic.

(c) Linkage concepts encountered in Psychical Research

When we turn to the linkage concepts encountered in Psychical research we find that they are essentially the same as those already reviewed. One of the commonest procedures reported is that of psychometry. This term was coined by an anthropologist Buchanan (11) to refer to the process whereby sensitives supposedly obtained their knowledge supernormally from an object held in their hands. Since the term is also

a technical one used by psychologists it is far from ideal, and Osty has proposed the term "Metagnomy" as a substitute. Despite its disadvantages the first term is the one more commonly used. It is proposed to examine the work of previous investigators of this technique.

There is no doubt that some impressive results have been obtained from psychometric procedures. Prince, (52) drew attention to one of these in a presidential address to the Society for Psychical Research. An unseen object was held between the hands of Mrs. King, a sensitive, and of the 34 testable statements she made, 33 were found to be correct, and the remaining one partially so.

One of the earliest investigators to give attention to the role played by the object psychometrised in mediumistic trance was Mrs. E.M. Sidgwick (75). In this investigation, Mrs. Piper, the medium, claimed to be producing messages from the surviving spirits of people who had died. She used articles once owned by these individuals to help her establish contact with them. Such articles were called "Influences".

The influences were supposed to bring the Control (the personality in control of the medium during trance) into touch with their present or former owners, and were sometimes spoken of as a source of information, that might be used to prove that the presence of an actual Sitter was not necessary. Articles that had been worn or handled by the owner were preferred, and were placed in the medium's hands, or near her writing block. Sometimes the medium held the articles throughout the sitting, sometimes only at intervals. Such influences were not held to be essential, but were claimed to be an aid in the process of getting messages.

Sidgwick reported that when the secondary personalities in control of the medium during trance were questioned, they made the following statements on the function of an influence.

It was claimed that an influence attracted its former owner, stimulated his associations, and made him feel at home. Owing to "its own light" the influence gave its owner strength to communicate, and helped him to "avoid being telepathically acted on by spirits". This was presumably what one of the controls (Rector) meant when he talked of the influence "holding the desired spirit apart from others who might unconsciously interrupt".

Another control (G.P.) claimed that the influence gave the control an impression of the spirit desired, enabling the control to pick out that spirit in the spirit world. In other instances it enabled the control to diagnose what was wrong with a distant patient. Once contact had been made, the influence was no longer necessary.

Two controls, (G.P. and Doctor) were sceptical of the possibility of being able to read the history of such an object from the object itself, without the intervention of a spirit.

It was claimed that the effectiveness of such influences diminished with the passage of time, that they deteriorated if carried over water, or if they were handled by many people, or if they came much in contact with other influences.

The effective part of an influence was described by controls as its "light" or "etherial".

Sidgwick comments that it would appear that the controls are themselves very vague as to how the influence enables contact to be made with living persons.

In commenting on such statements, Hodgson, in his first report (26) suggested that it might be the controls' belief in the efficacy of such articles that rendered them effective aids. He made the following suggestion:

"Now where the sitter knows the circumstances connected with the object, the associations will probably

form a particularly vivid cluster of experiences in his mind, conscious or subconscious, owing to the very presence of the object in his field of perception, and this may help (the control) to discover and disentangle the associations by direct thought-transference".

"Where on the other hand, the sitter is ignorant of the circumstances connected with the object, it may, at any rate, form a sort of point de repère enabling the control to get telepathically, through the mind of the sitter, at the mind of the distant living person."

Sidgwick, commenting on the statements made by the controls, says that the influences were apparently resorted to, and apparently found useful in Mrs. Piper's trances, in cases where the usefulness could not possibly have consisted of the things the trance personalities claimed, or in their being aids in telepathic selection. For example, she reported that Mrs. Piper would grasp a handkerchief known as "The Healer's Influence" to gain strength, though this handkerchief was totally unconnected with either the sitter or the control.

Mrs. Sidgwick consequently suggested that the use of influences might merely be part of a ritual, and that this ritual might have a steadying effect on the trance condition, helping perhaps to fix attention, and avert dream-like instability.

Another worker who has made considerable use of psychometric techniques is Dr. Osty. Unfortunately the writer has not been able to get direct access to his book "The Supernormal Faculties of Man", and has had to rely on quotations from Tyrrell's book "The Personality of Man" (93). Tyrrel quotes Osty's conclusions as follows:

"(1) After the sensitive, by holding the object, has once achieved psychological connection or rapport with the owner or contactor, the object may be destroyed without

affecting the sensitive's power to give information about him. Psychological rapport with the contactor once having been established, information is sometimes given about events which happened to him after the object has been destroyed.

(2) The material of which the object is made does not matter.

(3) If objects used in this way are allowed to touch one another, it does not make any difference.

(4) The length of time during which the owner has possessed the object or made contact with it does not matter.

(5) The lapse of time since the owner last touched the object does not matter.

(6) When once a sensitive has entered into the life of the owner of the object, the whole of that life is accessible and not merely the portion of it during which he possessed the object." (93, pages 176-7)

It is perhaps also worth noting that Osty was successful in getting one of his sensitives to make contact with a desired person through verbal direction only.

Another fact of interest is that Osty pointed out that his sensitives were by no means always correct, and that their errors were instructive. He found that they were not often affected by what people were consciously thinking or expecting, but that they were affected by their subconscious wishes, hopes, expectations and desires.

He also concluded that direct apprehension of clairvoyant knowledge was very rare in his sensitives. The knowledge that they showed came not through direct veridical hallucinations, corresponding to the objective reality, but rather through their own symbolic constructs.

Tyrrell is led to conclude from a study of Osty's findings that "it seems from these facts that the information acquired cannot in any sense be "locked up" in the object itself. The role of the object would seem to be rather to

canalise the sensitive's faculty and concentrate it in the right direction, though we have no information on how it does so. " (93, p 177)

Turning now to the work of Hettinger, Tyrrell quotes the following conclusions drawn from his first series of experiments.

"Neither the nature of the object 'psychometrised' nor the sentimental or emotional association therewith, if any, is the determining factor of the 'sensitive's' response." "The actual handling of the articles by the sensitive is not absolutely necessary for the exercise of the faculty." "Many of the items concern experiences by the owner after he has parted from the article." Tyrrell draws attention to the fact that these conclusions are in agreement with those of Dr. Osty. (93, page 129)

In "Exploring the Ultra Perceptive Faculty" (25) Hettinger surveys the findings from his second set of experiments, and makes the following points with reference to psychometry:

(1) "Distance tests without the use of an object belonging to a subject, unknown, or little known, to the Sensitive, failed to produce any results; hence the conclusion that, in order to establish a rapport between a Sensitive and a distant person such as referred to, the use of an article belonging to the Subject is absolutely necessary." (25, p.152)

(2) "If one Sensitive is psychometrising an article belonging to a Subject, and at exactly the same time an object of that Sensitive is being psychometrised by a second Sensitive (relay arrangement), we get the following possibilities:

(a) Both Sensitives may receive the same or similar impressions at exactly, or approximately exactly, the same time, or

(b) The first Sensitive may get one

impression of the picture contemplated by the Subject, and the second one get at the same time a different impression thereof, both being correct, or

(c) The first Sensitive may get no impression whatsoever corresponding to the picture, but the second one may do so." (25, p.155)

It will be remembered that the trance personalities of Mrs. Piper claimed that influences contained a mysterious "etheric" or "light" which rendered them effective. However Hodgson and Sidgwick both tended to reject the statements of the controls as being unreliable and vague, and both tended to attribute any effectiveness possessed by the influence as due to other psychological causes. The findings of Osty and Hettinger confirm this view. In some cases the articles are never touched by the sensitives, in other cases the article is only associated with the subject, and has never been handled by him, for example, a photograph, or a stone from the person's grave. Under such circumstances it is difficult to maintain the view that the object contains locked up in itself some special power or Mana, which the sensitive is able to utilise. We will therefore turn now to a study of theories based on the psychological effects of linkage, in particular that put forward by Whately Carington.

Carington's Association Theory of Linkage:

After publishing his experimental results, Carington followed them up with a theory that the telepathic process was essentially associative in nature. (14). Telepathy was made possible, he said, by the fact that certain common ideas existed in the minds of both agent and subjects. He called these common ideas "K-Ideas". He postulated that any association formed in the agent's mind with a K-idea, was also formed in the subjects' minds as well. Applying his theory to his own

experimental data, Carington suggested that the "Idea-of-the-Experiment" was the K-idea present in the minds of agent and subjects. When making his target drawings the agent tended to associate them with the "Idea-of-the-Experiment". When the subjects sat down to make their drawings, the "Idea-of-the-Experiment" was also uppermost in their minds as well, and so tended to bring with it the drawing that the agent had associated with it.

The crucial assumption that this theory makes is, of course, that any association made in the agent's mind with a K-idea is automatically made in other minds containing that K-idea as well.

Carington did not attempt to prove his theory, but rather attempted to show that it was not inconsistent with the observed facts, and that it was useful in explaining why they took the form they did. Thus if the telepathic process were essentially associative in nature, one would expect the sub-laws of association to operate: recency, frequency, vividness, similarity and contrast, nearness in space and time.

Carington showed that his data did conform to some of these sub-laws. His subjects scored the greatest proportion of hits on a given target at about the time it was actually being exposed, with a gradual precognitive and postcognitive falling off. He claimed that this was what would be expected if the law of recency of association were operating.

He pointed out that under the law of frequency, it was to be expected that if repetition of an associative act, regardless of the mind in which it takes place, tends to strengthen the association in the telepathic process, then the greatest number of hits should be scored on target material which has been associated in the greatest number of minds. In one of his experiments Carington made out a list of 216 names of possible target originals. This list was sent to

another person, who selected 50 titles and drew illustrations of them. These were placed in opaque envelopes, and sent to a third person, who together with Carington, randomly selected 10 which were to be used as originals. Carington found that his subjects tended to score hits on all the objects listed in proportion to the number of people who had handled them, thus:

Originals listed, drawn & used	21.2 av. score.
Listed and drawn, but not used	4.2
Listed only	2.2
Controls	0.2

Commenting on the light shed by his theory on the linkage problem, Carington says:

"One of the difficulties of the subject has always been to understand how it is that the subject contrives to pick out the experimenter's (or other "sender's") thoughts, etc., from the welter of assorted images, etc., which must be supposed available to him; and mutatis mutandis, of course, if he is supposed to be "seeing" or "sensing" the object. To deal with this it has been usual to assume some kind of 'rapport' or 'affinity' or the like, or something akin to 'resonance' between sender and recipient, though no kind of mechanism capable of explaining it has been suggested, as far as I know."

"The Association Theory does not need any special mechanism of this kind; it affirms that the required idea is not, properly speaking, 'selected' at all, but is given, so to say, a better chance of coming into the subject's mind by virtue of the fact that it has been associated by the experimenter with the idea of the experiment, E, whereas other ideas have not been so associated, and therefore lack this advantage." (14, page 67)

Carington goes on to examine the idea of rapport, and suggests that the more K ideas the agent and subject have in common, the greater the likelihood of telepathy between them. On this hypothesis it would be expected that the subjects who

had been supplied with a K idea in the form of a photograph would do better than those who had not. He found that his data did in fact show this difference, although the experimental design was not such that the difference could be certainly attributed only to the influence of the photograph. He suggested finally that an experiment to test the effectiveness of artificially introduced K ideas would be very useful.

Although Carington's theory created quite a lot of interest in parapsychological circles, (Reeves for instance published a review in the Journal of the American Society for Psychical Research, (53)) it appeared to be open to some of the criticisms which had been levelled against the old Associationist school of Psychology, and tended to be discounted on philosophical rather than empirical grounds. His work stimulated many to try and repeat his "repeatable experiment" (E.g. 1; 59; 66; 82; 83; 84; 85; 86; 90; 91 & 98), and revived an interest in the use of drawings as target material, but none attempted to test his theory of K ideas empirically. Though the work of these writers was interesting and suggestive, the present writer was left with the feeling that they were incidental observations, made from experiments not really designed to investigate linkage as such. There appeared to be a great need for a systematic experimental investigation of linkage in its own right. It was decided to make a beginning with this task.

The logical starting point seemed to be to establish whether or not different types of linkage material had any effect on the E.S.P. scoring level of subjects: that is, whether subjects supplied with linkage material did better than an equated group of subjects working without such linkage. It was decided to test the effectiveness of three types of linkage material: personal object links, K idea links, and

symbolic descriptive links.

In the event of any of these types of linkage being found effective, a second important question would arise, namely, on what qualities of the links did the effectiveness depend? Were the links effective merely because of their psychological suggestion value, or were other factors also involved? What was the relative effectiveness of the different types of linkage? It was decided to design an experimental procedure that would produce information on these problems.

(d) Review of the Experimental Procedure:

It will be remembered that the subjects were divided into two groups --- an Experimental Group, and a Control Group. The subjects in the Experimental Group were given correct linkage material, whereas the subjects in the Control Group were given completely false linkage material, though they were led to believe that their linkage material was correct.

For the first week of the experiment (Section "A") neither Experimental nor Control Group Subjects were provided with any links with the agent. This section was intended to establish whether the two groups of subjects were initially equal in hit-scoring ability, and also to establish what frequency of hits would be obtained with no linkage operating.

In the second week, (Section "B") links were introduced in the form of squares of colour, with the aim of testing the K idea theory. The agent each day associated his drawings with a particular colour, and the Experimental Group Subjects were provided with an identical square of colour on their reproduction sheets. On Carington's hypothesis the provision of these links should cause more hits to be scored in section B than were scored in section A of the experiment. Also, more hits should be scored on the actual

day the target was exposed than on other days, and in other sections of the experiment, since different K ideas specific for each day were provided. The Control Group Subjects were also given squares of colour, but these differed from the colours actually being used by the agent.

The third week, (Section "C") was designed to test the value of providing the subjects with a personal article from the agent they were trying to contact. The Experimental Group Subjects were each given a square of handkerchief which had been used by the agent for some months. The Control Group Subjects were given handkerchiefs of the same size, but which had never come in contact with the agent at all. This section was of course designed to test the value of psychometry.

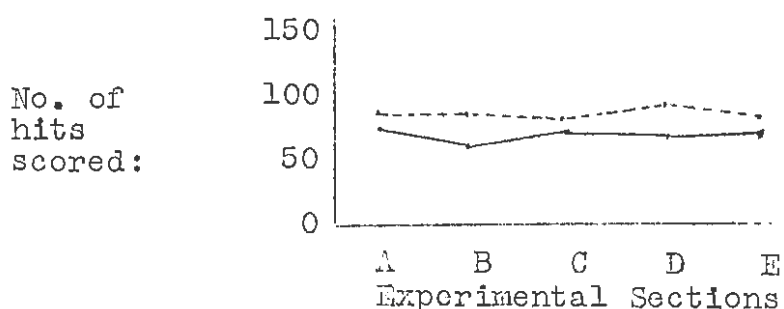
For the fourth week (Section "D") the subjects were provided with symbolic, verbal links, in the form of a photograph of the agent, and a personal description of him. The Experimental Group Subjects were given correct photographs and descriptive material; the Control Group Subjects were given the photograph of a man unconnected with the experiment, and personality material relating to someone else who had taken no part in the experiment. It was thought that this section would provide an additional test of Jarrington's theory, and of the effect of emotional rapport between subjects and agent.

Finally in the fifth week (Section "E") the same linkage material used in section D was again issued to the subjects. The agent, however, again associated each pair of drawings with a specific square of colour, but this time no squares of colour appeared on the subjects' books. (It had originally been intended to put such squares of colour on their reproduction sheets, as was done in section B, but the complaints made by subjects after section B that the colours had been a hindrance, led the writer to eliminate them from the last section.

The Linkage material can thus be summarised as follows:

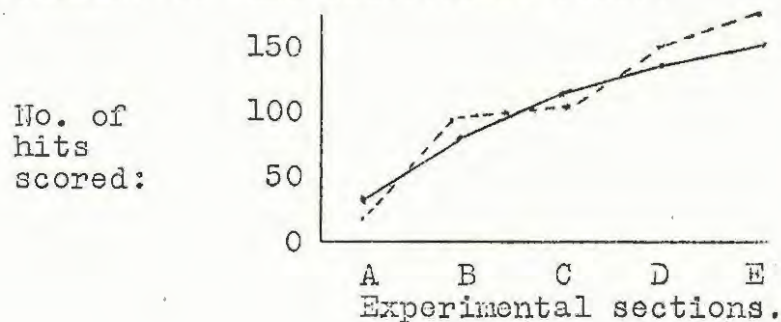
<u>Section</u>	<u>Linkage Material</u>	<u>Problem investigated:</u>
A	Nil	(a) Equality of groups (b) Normal scoring rate without linkage.
B	Squares of Colour	(c) Effectiveness of colours as K ideas.
C	Squares of Handkerchief.	(d) Effectiveness of personal articles as psychometric links.
D	Photograph & personal description of agent.	(e) Value of symbolic links (f) Effect of emotional rapport (g) The effect of providing a large no. of potential K ideas.
E	Photograph & personal description of agent & Squares of colour used by agent only	(h) As in section D, but plus the agent's associations of his drawings with a particular colour.

It was thought that this arrangement of the experimental procedure would allow the data to be interpreted as follows. If all the linkage material provided was ineffective, then neither the Experimental Group Subjects nor the Control Group Subjects should show any improvement on the scoring rates established in section A of the experiment. Illustrated graphically, such results would take the approximate form of two straight lines:

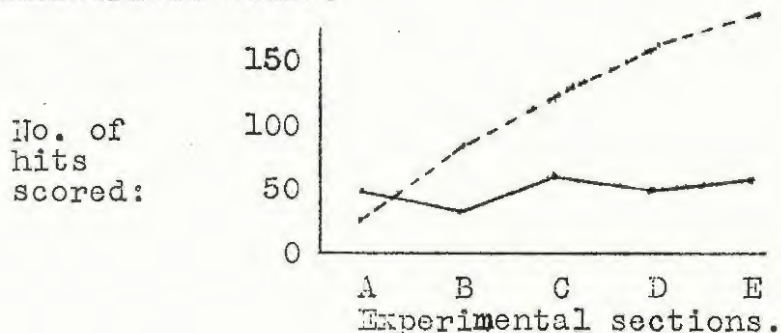


However, if both Control and Experimental Group Subjects showed an equal improvement on their score levels established in section A, then it could be argued that this improvement

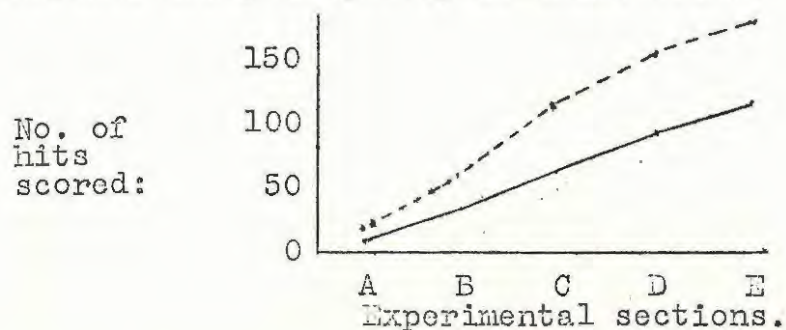
was due either to uncontrolled factors influencing both groups of subjects, or that it was due to the suggestion value of the linkage material introduced, or to a combination of these two factors. Depicted graphically, such results would take a form similar to that illustrated below:



If, however, correctness of the linkage material had some virtue in itself, other than its mere suggestion value, then the Experimental Group Subjects performance should show an improvement on the rate established in week A, whereas the Control Group Subjects should show no significant improvement as is illustrated below:



Finally, if the linkage material was effective both because of its suggestion value, and because of other qualities such as its correctness, it would be expected that both Control and Experimental Group Subjects would show an improvement on their scoring rate established in week A, but the improvement of the Experimental Group Subjects should be greater than that of the Control Group Subjects, as shown below:



It is possible to conceive of factors other than the linkage material producing these effects. For instance, if for any reason the associations of the subjects were less free in the first part of the experiment than in subsequent parts, they would stand less chance of scoring hits initially than later. If the experiment had been started in examination time, it might be argued that at first the subjects' guesses were determined largely by the unconscious complex of ideas associated with examinations. It was thus difficult for them to make genuinely free guesses. Later, as the press of examinations receded, their associations became less "tied" and the probability of their hitting one of the rare originals became correspondingly greater, even on a purely chance basis.

It could also be argued that a subject whose associations are "tied" in this way offers a much greater resistance to any E.S.P. impression getting through than one whose associations are more dissociated and fluid.

Such general factors would be expected to cause an increase on the hits scored on the Control Series Originals as well as on the Experimental Series Originals used as targets in the actual experiment. Consequently, by comparing the scoring rate on the Experimental Series Originals with the rate on the Control Series Originals, it will be possible to determine whether any improvement is genuinely due to the effect of linkage, or whether it is merely due to the operation of uncontrolled general factors.

Consequently, in the examination of the data that follows it will be necessary to consider hits scored on both Experimental and Control Series Originals by both Experimental and Control Group Subjects. The first step will be to examine the data to find whether it shows any changes in scoring rate from week to week, and if inspection does reveal such changes, then the next step will be to assess the significance of these.

(2) Graphical presentation of the raw scores:(a) Adjustment of the raw scores:

Before a direct comparison can be made, the scores must be adjusted to allow for the varying number of drawings returned in the different sections of the experiment by the different groups of subjects. To enable such a comparison to be made, the raw score frequencies have been converted to the frequencies that would be expected if each group of subjects had returned 2,000 drawings in each section of the experiment. Thus in section A of the experiment the Control Group Subjects returned 1,801 drawings, scoring 254 hits, and the Experimental Group Subjects returned 1,741 drawings, scoring 244 hits. If the Control Group Subjects had returned 2,000 drawings instead of 1,801, their expected number of hits would be

$$\frac{2,000}{1,801} \times 254 = \underline{282.1 \text{ hits.}}$$

In the same way, the expected number of hits for the Experimental Group Subjects would be:

$$\frac{2,000}{1,741} \times 244 = \underline{280.3 \text{ hits.}}$$

These adjusted frequencies have been used for all the graphs that follow:

(b) Graphs:

GRAPH No: 4

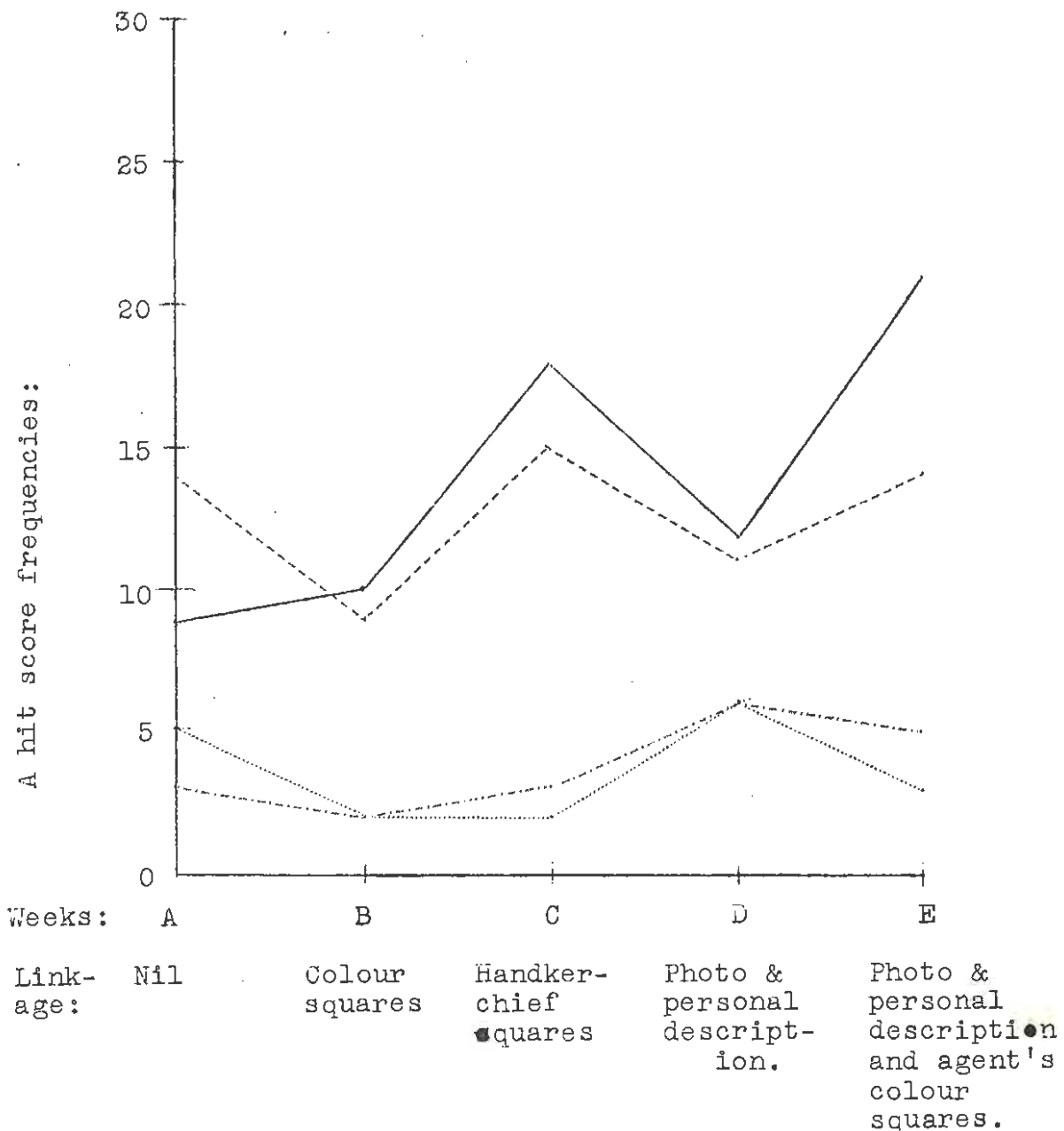
EFFECT OF LINKAGE ON THE RAW SCORING RATES:

EXPECTED SCORES PER 2,000 TRIALS

PER SUBJECT GROUP PER WEEK.
(Data Sheet No.: 17)

<u>Legend:</u>	
<u>Hits on Experimental Series Originals:</u>	<u>Hits on Control Series Originals:</u>
By Experimental Group Subjects: ———	By Experimental Group Subjects: - - - - -
By Control Group Subjects: - - - - -	By Control Group Subjects: ·····

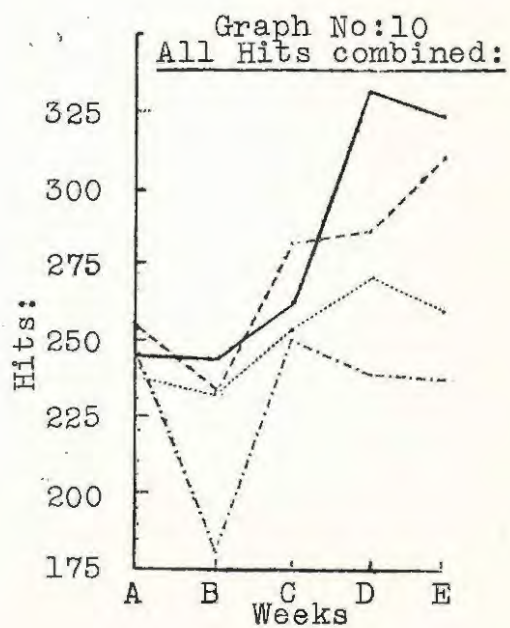
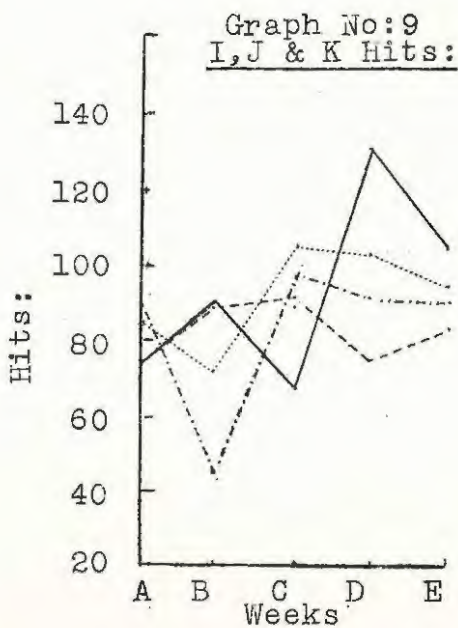
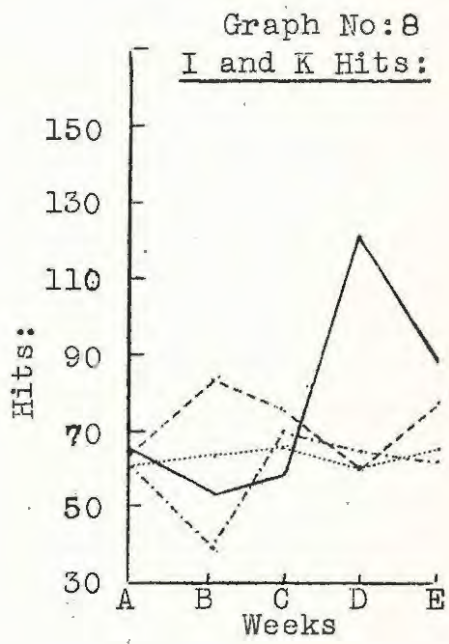
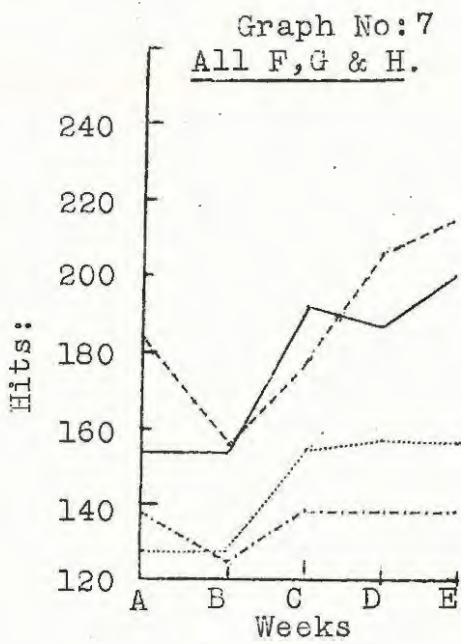
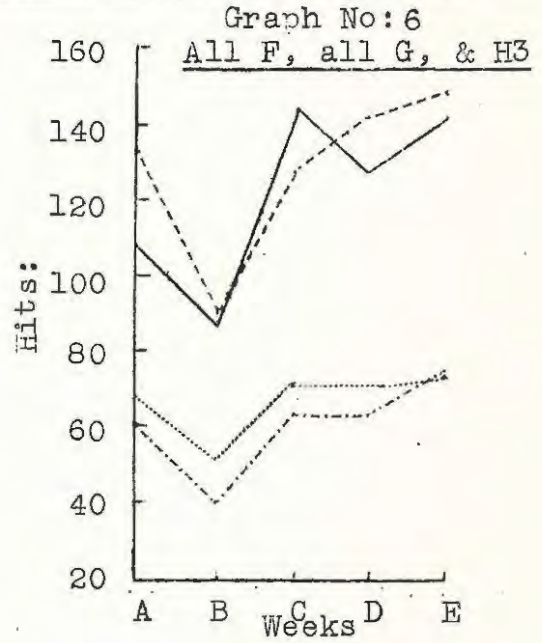
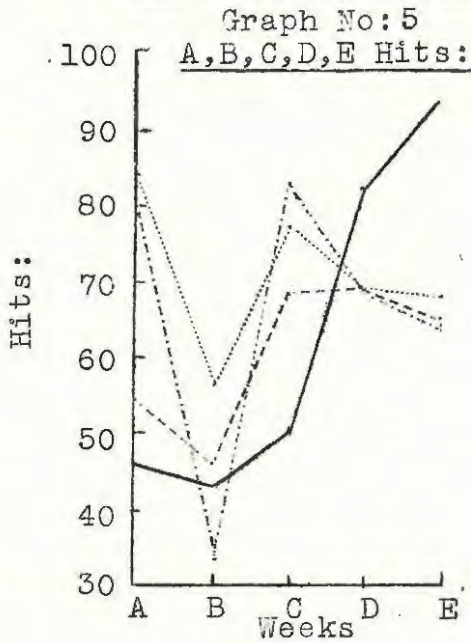
UNCORRECTED A HITS:



EFFECT OF LINKAGE ON THE RAW SCORING RATES:
EXPECTED SCORES PER 2,000 TRIALS
PER SUBJECT GROUP PER WEEK.

(Data Sheet No.: 17).

Legend:	
<u>Hits on Experimental</u>	<u>Hits on Control</u>
<u>Series Originals:</u>	<u>Series Originals:</u>
By Experimental	By Experimental
Group Subjects: ———	Group Subjects: - - - - -
By Control	By Control
Group Subjects: - - - - -	Group Subjects: - - - - -

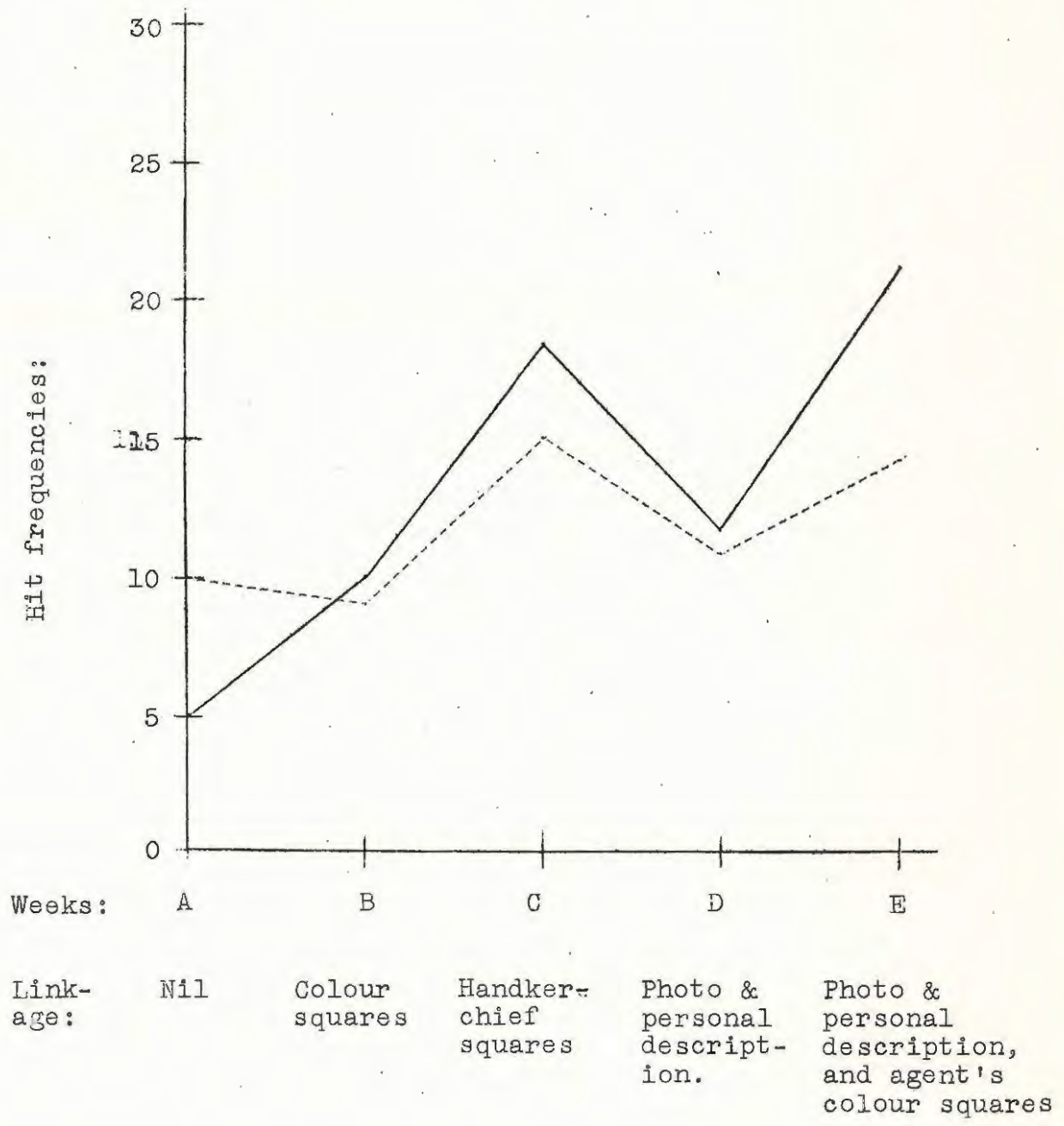


GRAPH No: 11

EFFECT OF LINKAGE ON SCORING RATES CORRECTED
FOR BUNCHING ERRORS.
EXPECTED SCORES PER 2,000 TRIALS
PER SUBJECT GROUP PER WEEK.
(HITS ON EXPERIMENTAL SERIES ORIGINALS ONLY)

Legend:
 Experimental Group Subjects: ———
 Control Group Subjects: - - - - -

A HITS CORRECTED FOR BUNCHING
 (From Data Sheet No.; 18)

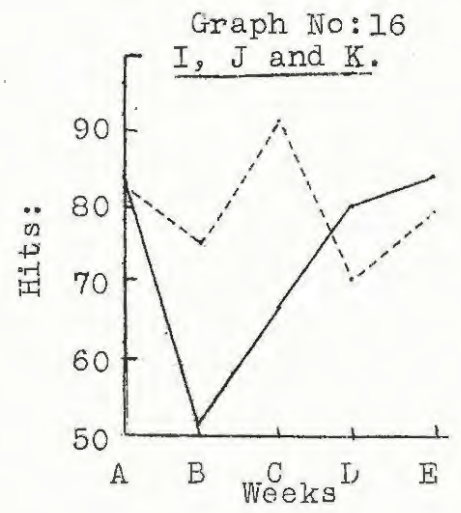
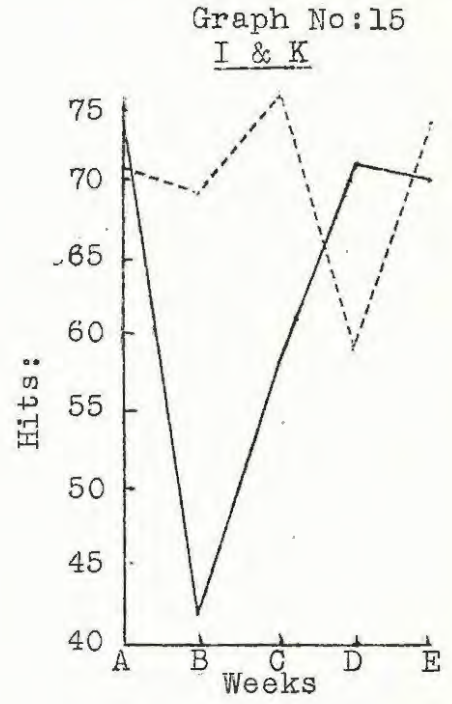
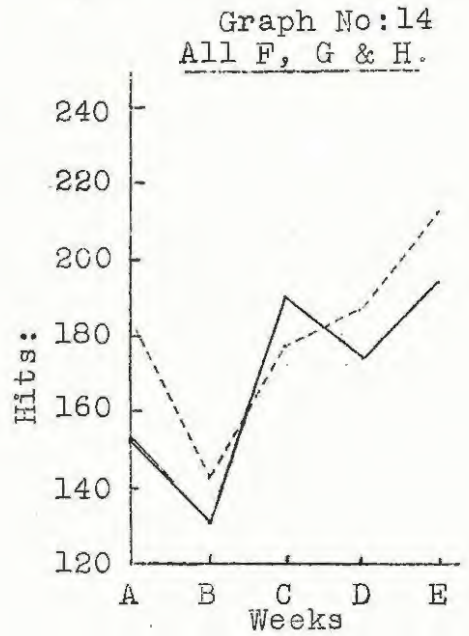
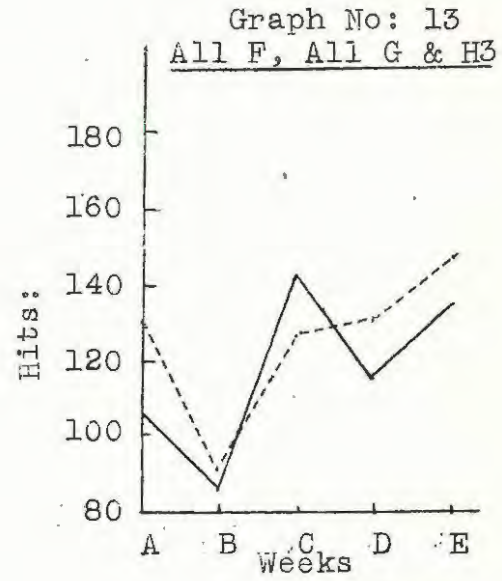
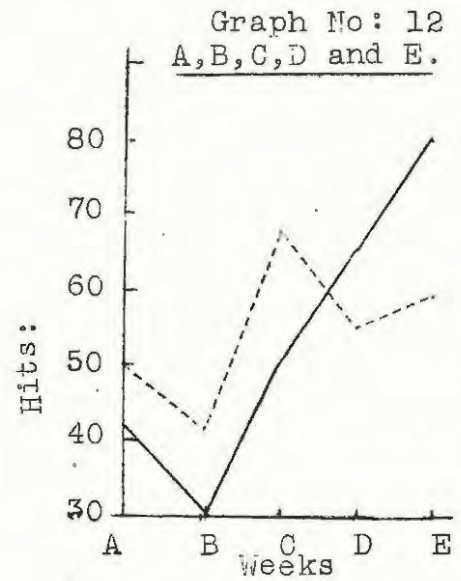


Checked? Yes.

GRAPHS Nos:12 to 16

EFFECT OF LINKAGE ON SCORING RATES CORRECTED FOR BUNCHING ERRORS
EXPECTED SCORES PER 2,000 TRIALS PER SUBJECT GROUP PER WEEK
(HITS ON EXPERIMENTAL SERIES ORIGINALS ONLY)

Legend:
 Experimental Group Subjects: ———
 Control Group Subjects: - - - - -



Data Taken from Source Data Sheet 18.

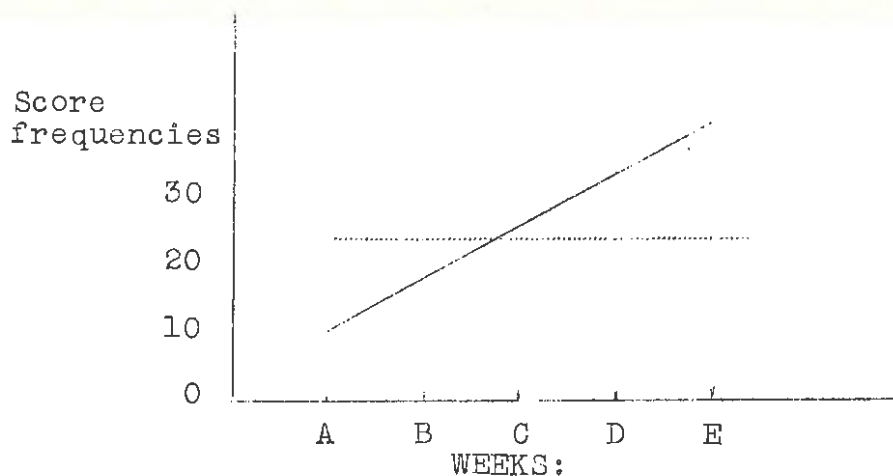
N.B. "All hits Combined" showed no marked bunching and so were not corrected.

Checked ? Yes.

(c) Trends in scoring rate revealed by an inspection of the graphs.

Let us first consider graph 4 showing fluctuations in Raw A hits. Since the component originals in both series were equated in difficulty in terms of A hits, it is permissible to compare the scoring rates on these two series of originals, as well as compare the scores returned by each group of subjects on each series. The graph shows clearly that both groups of subjects score more hits on the targets used in the experiment than on the control originals. The two top curves, depicting hits on Experimental Series Originals, at first inspection seem to show quite similar variations, and to follow one another closely, suggesting that similar factors were operating to produce them. In terms of the experimental design these could either be suggestion effects due to the linkage material, or might be due to other uncontrolled factors. Such uncontrolled factors would be expected to affect hits on the control originals as well as the targets, and since the bottom two graphs, depicting hits on C.S.O. do not follow the pattern set by the hits on the E.S.O., it would seem likely that suggestion rather than uncontrolled factors is causing the similarity of the top two graphs.

Although the hits on the E.S.O. of the E.G.S. and C.G.S. thus appear to fluctuate in unison, there is an important difference. Inspection will show that the E.G.S. scoring rate (solid line) starts low, in week A, and then improves, so that it shows an upward trend, finishing strongly in week E. If a smoothed straight line were fitted to this graph it would show a steady rise from week A to week E. By contrast, the hits scored on the E.S.O. by the C.G.S. start high in week A, but fluctuate about this point, ending at about the same level. A smoothed line fitted to this data would lie almost horizontally, as shown in the sketch overleaf:



Reference to Graph 11 depicting A hits corrected for bunching shows that the bunching correction does not appreciably alter these trends.

Turning to a consideration of the other types of hit presented in Graphs 5 to 11 a more complex picture confronts us. Here too, however, examination reveals certain well-marked tendencies. If we turn our attention first to the solid line depicting E.G.S. hits on E.S.O. it will be seen that in every case the E.G.S. scored considerably more hits in weeks D and E than they did in week A, thus confirming the trend shown by the A hits. In some instances they also show an improvement in week C, but this is neither so marked, nor so consistent. By contrast the scoring rates of the C.G.S. on the E.S.O., (shown by the dashed line -----), do not disclose such a consistent trend, weeks D and E producing sometimes more, sometimes less hits than in week A.

It will be remembered that when the subjects were issued with squares of colour as linkage material, in week B, some complained that they found the colours hindered them in getting impressions for their drawings, by constraining their associations. Interestingly enough, the graphs support this contention. There is a general falling off in week B, not only with hits on E.S.O., (lines — and -----), but also in the hits scored on the C.S.O., (lines and -----). Another point of interest is that in general, this falling off is least marked in the case of the E.G.S. hits on C.S.O. (solid line).

Graphs 12 to 16 depicting the scores corrected for bunching, bring out clearly the drop in week B, but the corrections have tended to confuse the other relationships commented on, in hit types I and K, and I, J and K (Graphs 15 and 16).

Let us next examine the data showing hits on the C.S.O., (----- for the E.G.S., and for the C.G.S.). Here the picture confronting us is much more random, and with the exception of the drop in week B, no consistent trends are obvious.

It would thus seem that the data tends to show interesting effects that would repay further investigation. On common-sense grounds it would seem likely that at least some of the linkage material had had an effect, but it should be stressed that although the above trends are suggested by an inspection of the graphs, we need to know whether they are large enough to reach statistical significance before we can draw any conclusions from them. The next step will therefore be to test the statistical significance of the differences observed.

(3) General discussion on the choice of suitable statistical techniques for the Linkage assessments.

The task confronting us is two-fold, firstly, to find whether the fluctuations shown are large enough to be significant, and in the event of them proving to be so, secondly, to find whether they can be attributed to the introduction of Linkage Material, or merely to other uncontrolled factors.

Four possible types of measure are available for consideration. These are: the Raw Scores and the Scores Corrected for Bunching, and the Expected frequencies per 2,000 trials for each of these. Although the Expected Frequencies were used to construct the graphs, these are not suitable, for

in many techniques of statistical assessment the significance of any difference can be increased by increasing the number of trials. The actual trials returned in the experiment ranged from 1,678 to 1,801 per group per week, whereas the Expected Scores are based on a standard 2,000 trials per week. Consequently any significances calculated from them would slightly overestimate the significance of the actually observed differences. We are left with the Raw Scores, and the Scores Corrected for Bunching. Of these two types the Raw Scores are perhaps preferable, for the reasons given below.

There is no reason that the writer can see, why uncontrolled factors producing spurious hits should favour the E.G.S. or the E.S.O. more than their control counterparts. There are moreover 50 originals in each series, giving a fair chance for atypical originals in one series to be balanced by similar originals in the other series. A critic might argue, however, that although there was no reason for chance to favour the E.G.S. hits on E.S.O. it just did. Thus a power failure in a residence of subjects necessitated the use of candles: this in turn caused these subjects to draw more candlesticks (one of the originals), and by chance the subjects happened to be E.G.S. and by chance again the original happened to form part of the E.S.O. Alternatively, it might be argued that the Linkage Material itself, which was issued to the E.G.S. was not completely equated with that issued to the C.G.S. For example, the material issued to the E.G.S. perhaps mentioned that he was fond of music, a violin happened to be one of the target originals, and thus a spurious burst of hits on violin resulted. Undoubtedly such happenings did occur, but inspection of the original hit distributions shows that they occurred equally for each subject group and series of originals, and that they were not confined merely to the E.G.S. hits on E.S.O. The onus would thus appear to be on any critic to show

specifically why such uncontrolled factors should favour the E.G.S hits on E.S.O., and not the other classes of data.

The Scores Corrected for Bunching attempt to reduce any a-typical spurts of hits to the average for other weeks, and some readers may prefer to base their conclusions on them. The writer has thus carried out calculations for significance on this type of hit as well. He feels, however, that these corrections may have eliminated some genuine hits produced by E.S.P. So little is known of the way in which E.S.P. functions that we cannot say that it does not on occasions produce hits in spurts. Suppose that the telepathic message is received on a subconscious level, and needs some sort of associative link to help it come to conscious awareness. Suppose too, that among the E.S.O. is "Father Christmas", and among the C.S.O. "Mr Chad". During the course of the experiment some subjects go to a lecture on "Christmas Customs", and subsequently there is a spurt of hits on "Father Christmas". By contrast, no spurt of hits on "Mr Chad" is produced, even when posters prominently depicting him are displayed in various places frequented by the subjects. The experimental data when inspected seems to show this kind of effect. The writer would like to suggest that an E.S.P. impression dormant in the subconscious mind of the subject renders him responsive to trigger stimuli in the environment. Thus because "Father Christmas" is a target original, the subjects respond to any Christmas stimuli in the environment, and react to them by drawing Father Christmasses. Because "Mr Chad" is not a target original, posters of him do not stimulate the subjects to draw him, and no false spurt of hits is produced.

The writer also feels that the corrections applied for bunching errors were admittedly crude, and to some extent an interference with the distributions produced by random chance factors. For this reason, and the one outlined above, he personally prefers to base his conclusions on the Raw Scores. The Scores Corrected for Bunching have also been assessed for significance in case any reader would prefer to use them.

The writer consulted a number of authorities in order to find what type of statistical test would be suitable for the data. The suggestions ranged all the way from those who advocated the application of regression techniques, to one authority who thought that the trends shown by the data were so clear that they did not need statistical assessment. In the end the writer was still faced with the task of choosing between the various courses advocated.

The use of regression techniques would reveal whether or not the fluctuations shown over the five weekly sections of the experiment taken as a whole were large enough to be significant. It would also be possible by these means to find mathematically what type of curve best represented the experimental results obtained, and hence the mathematical equation and number of factors responsible for producing the results. The writer would welcome any reader carrying out such an assessment on the data, but feels that it is beyond his personal statistical competence to effect. Rightly or wrongly he feels that the main interest in the data lies in the changes in scoring rate between week A, when no linkage material was given, and other weeks, when a different type of linkage was introduced each week. This led him to doubt the validity of assessing the experiment as a whole over all five weeks combined. He decided to adopt the simplest procedures possible which would give valid results.

Initially the writer carried out an assessment in which he took the score returned in week A as the theoretically expected score, and assessed each subsequent week against this in order to find whether its score differed significantly. Luckily this method was submitted to Dr. J.G. Pratt for approval, who in correspondence pointed out that it was open to serious objections:

".....I am afraid I would protest your using week A as standard, inasmuch as it assumes that by some magic A gave

exactly the theoretical expectation and therefore no question of the variance of A should arise. Why not instead work out the chi square on the following scheme:

	Week A	Week B	Week C	etc	Total
Hits					
Misses					
Total					

In that way you would take the total results as giving you the best estimate of chance, and each week's results would be used in getting a measure of the variation among the weeks. Thus you would have a 2 by x table, x being the number of weeks, and d.f. being $x - 1$. " (Letter dated September 26th, 1955.)

The method described above by Dr. Pratt has been adopted for the first assessment of the linkage data, which follows. In a second assessment the fluctuations between week A and each other week have been tested, using a modified procedure designed to meet Dr. Pratt's objections to the discarded method. Instead of taking the scores of week A as the theoretically expected frequency, the scores made in week A and week X are totalled, and the average (allowing for the varying number of trials in each week) taken as the theoretically expected frequency. Each of these methods will be explained fully before it is used.

(4) Assessment of the Significance of the Linkage Scores
fluctuations by averaging the hits from all weeks pooled
to obtain theoretically expected frequencies for a χ^2 test.

(a) Statistical Method:

As an example of the statistical method let us take calculation C 104, dealing with Raw A type hits made by E.G.S. on E.S.O
 The basic data is as follows:

	<u>Week A</u>	<u>Week B</u>	<u>Week C</u>	<u>Week D</u>	<u>Week E</u>
Hits:	9	10	18	12	21
Trials:	1741	1760	1702	1702	1678

If no significant fluctuations have taken place, we would expect to find the hits distributed between these weeks in proportion to the number of trials made in each week. Hence the first step is to find the total number of hits, = 70; and the total number of trials, = 8583. Next we find the theoretically expected hits by multiplying the total hits by the trials made in the week, divided by the total trials. Thus in week A one would expect $\frac{1741}{8583}$ or .20284 of the total hits to be found. This gives $70 \times .20284 = 14.20$ (rounded off correctly to two decimal places) as the theoretically expected hits for week A. Subtracting this from the trials, $1741 - 14.20$ gives the expected misses = 1726.8 for week A. Subtracting the observed hits from the expected hits gives the deviation "D" which will be the same for misses as well as hits, = 5.20. This is squared, = 27.04, and divided by the expected hits 14.2, to give the χ^2 value for the week A hits = 1.90. D^2 is also divided by the expected misses to get a χ^2 value for the misses. This procedure is repeated to get values for weeks B, C, D and E. (The theoretically expected frequency for the final cell has been found by calculation, and not subtraction from the marginal totals. Because it is necessary to round off, totals for the theoretically expected frequencies occasionally deviate plus or minus .01 from the total hits observed, in a few of the calculations). The χ^2 values for hits and misses are then totalled, giving 8.66 for 4 degrees of freedom, and entering appropriate tables of χ^2 shows P (two tail) to lie between .10 and .05. This is approaching significance, but does not reach the .02 level which we are going to set as our level of confidence in these assessments. We have thus insufficient grounds for rejecting our null hypothesis, and little justification for assuming that anything but chance is producing the fluctuations shown by these scores.

All P values given for these calculations, and shown in the summary tables are for a two-tailed test of significance, since this statistical method does not take into account the direction of the deviations, and since at the moment we are merely interested in finding out whether any fluctuations are significant, irrespective of the direction of the deviation from chance .

In these calculations only Raw Scores are assessed.

(b) Calculations C 104 to C131:

Table XIV:

Linkage Effects: Score Fluctuations for Experimental Group
Subjects for all weeks pooled tested for significance.

Calculation
 Serial No.: C 104.

Source Data
 Serial Nos.: D 1. & 16.

Raw A hits on Series E Originals.

Total trials: 8,583.00

Total hits: 70.00

Factor:	Week A	Week B	Week C	Week D	Week E
	.20284	.20506	.19830	.19830	.19550
Trials:	1741.0000	1760.0000	1702.0000	1702.0000	1678.0000
Hits Expected	14.20	14.35	13.88	13.88	13.69
Hits Scored	9	10	18	12	21
D	5.20	4.35	4.12	1.88	7.31
D ²	27.04	18.9225	16.9744	3.5344	53.4361
$\frac{D^2}{\text{Exp. Hits}}$	1.90	1.32	1.22	.25	3.90
Misses Expected	1726.8	1745.65	1688.12	1688.12	1664.31
$\frac{D^2}{\text{Exp. Miss}}$.02	.01	.01	.00	.03

χ^2 totals: Hits: 8.59 + Misses: .07 = 8.66 (for 4 d.f.)

P two tail lies between .10 and .05

Calculation
 checked ? Yes.

TABLE XIV.
LINKAGE EFFECTS: SIGNIFICANCE OF RAW SCORE FLUCTUATIONS
FOR ALL WEEKS POOLED.

Raw Hits made by Experimental Group Subjects on		
Hit type:	(A) E.S.O. (Calculations C 104 to C 110).	(B) C.S.O. (calculations C 111 to C 117).
	P (two tail) is less than:	P (two tail) is less than:
A	.10	(too small)
A,B,C,D and E	<u>.001</u>	<u>.001</u>
All F, all G, H3	<u>.001</u>	<u>.02</u>
All F, G and H	<u>.01</u>	.70
I and K	<u>.001</u>	.05
I, J and K	<u>.001</u>	<u>.001</u>
All combined	<u>.001</u>	<u>.001</u>

Raw Hits made by Control Group Subjects on:		
Hit type:	(C) E.S.O. (Calculations C 118 to C 124).	(D) C.S.O. (Calculations C 125 to C 131).
	P (two tail) is less than:	P (two tail) is less than:
A	.80	(too small)
A,B,C,D and E	.20	.20
All F, all G, H3	<u>.01</u>	.50
All F, G and H	<u>.02</u>	.10
I and K	.20	.98
I, J and K	.70	.10
All combined	<u>.01</u>	.30

N.B. All differences reaching the .02 level of confidence have been underlined.

Transcription checked ?
Yes.

(c) Conclusions drawn from Table XIV: "Significance of the Raw Score fluctuations for all weeks pooled."

Reference to the above table shows that the most consistent and significant fluctuations in score have occurred in the hits on E.S.O. made by E.G.S. However, since significant fluctuations also occur in the hits on C.S.O. by E.G.S., and since the C.G.S. also show a number of significant fluctuations, we are not justified in concluding that these effects were due to the Linkage Material operating by linking the subjects to the agent and making it easier for them to score hits on the E.S.O. The results point to the operation of some general factor common to blocks (A), (B) and (C) of the data. It will be remembered that the subjects complained that the squares of colour issued to them in week B constricted their associations, and made it more difficult for them to think of responses. The graphs also showed a fairly consistent falling off in week B. It may be that some of the significant fluctuations found in blocks (B) and (C) of the data are being caused by the hindrance effect exerted by the squares of colour. Fortunately this hypothesis can be tested quite simply by eliminating week B, and carrying out the same type of assessment on weeks A, C, D and E alone. This has been done in the calculations which follow.

(5) Assessment of the Significance of the Linkage Score Fluctuations by eliminating week B, and averaging the hits in weeks A, C, D and E to obtain theoretically expected frequencies for a χ^2 test.

(a) Statistical Method: This is similar to that used in calculations C 104 to C 131, except that week B is not included, and the final χ^2 total is for 3 d.f.

The calculations are carried out both on the Raw Scores, and the Scores Corrected for Lunching.

(b) Calculations C 132 to C 159

Tables XV and XVI

Linkage Effects: Score Fluctuations for Experimental Group
Subjects for weeks A C D & E pooled tested for significance.

Calculation

Source data

Serial No.: C 132.

Serial Nos.: D 1 & 16.

Raw A hits on Experimental Series Originals.

Total Trials: 6,823.00

Total hits: 60.00

Factors	Week A .25517	Week C .24945	Week D .24945	Week E .24593
Trials	1741.0000	1702.0000	1702.0000	1678.0000
Hits Expected	15.31	14.97	14.97	14.75
Hits Scored	9	18	12	21
D	6.31	3.03	2.97	6.25
D ²	39.8161	9.1809	8.8209	39.0625
$\frac{D^2}{\text{Exp. Hits}}$	2.60	.61	.59	2.65
Misses Expected	1725.69	1687.03	1687.03	1663.25
$\frac{D^2}{\text{E. Misses}}$.02	.01	.01	.02

χ^2 totals: Hits: 6.45 + Misses: .06 = 6.51 (for 3 d.f.)

P two tail lies between .10 and .05

Calculation
checked ? Yes.

TABLE XV.

LINKAGE EFFECTS: SIGNIFICANCE OF SCORE FLUCTUATIONS FOR
WEEKS A, C, D AND E POOLED, WITH WEEK B
ELIMINATED.

Raw hits made by Experimental Group Subjects:		
Hit type:	(A) On E.S.O. (Calculations C 132 to C 138)	(B) On C.S.O. (Calculations C 139 to C 145)
	P (two tail) is less than:	P (two tail) is less than:
A	.10	(too small)
A,B,C,D and E	<u>.001</u>	.50
All F, all G, H3	.05	.50
All F, G and H	<u>.02</u>	.99
I and K	<u>.001</u>	.80
I, J and K	<u>.001</u>	.90
All combined	<u>.001</u>	.98

Raw hits made by Control Group Subjects		
Hit type:	(C) On E.S.O. (Calculations C 146 to C 152).	(D) On C.S.O. (Calculations C 153 to C 159).
	P (two tail) is less than:	P (two tail) is less than:
A	.90	(too small)
A,B,C,D and E	.50	.70
All F, all G, H3	.70	.95
All F, G and H	.20	.20
I and K	.30	.90
I, J and K	.50	.50
All combined	.05	.30

N.B. All differences reaching the .02 level of confidence have been underlined.

Transcription checked ?
Yes.

TABLE XVI.

COMPARISON OF THE SIGNIFICANCE OF RAW SCORE FLUCTUATIONS
WITH SCORES CORRECTED FOR BUNCHING CALCULATED ON WEEKS A,
C, D and E POOLED, BUT ELIMINATING WEEK B.

(HITS ON EXPERIMENTAL SERIES ORIGINALS ONLY.)

Hits by Experimental Group Subjects on E.S.O.		
Hit type:	(A) Raw Scores: (Calculations C 132 to C 138) P (two tail) is less than:	(B) Scores Corrected for Bunching (Calculations C 160 to C 165) P (two tail) is less than:
A	.10	<u>.01</u>
A,B,C,D and E	<u>.001</u>	<u>.001</u>
All F, all G, H3	.05	.05
All F, G, and H	<u>.02</u>	.05
I and K	<u>.001</u>	.70
I J and K	<u>.001</u>	.50
All combined	<u>.001</u>	(not corrected)

Hits by Control Group Subjects on E.S.O.		
Hit type:	(C) Raw Scores: (Calculations C 146 to C 152) P (two tail) is less than:	(D) Scores Corrected for Bunching (Calculations C.166 to C 171) P (two tail) is less than:
A	.90	.70
A,B,C,D and E	.50	.50
All F, all G, H3	.70	.70
All F, G and H	.20	.30
I and K	.30	.30
I, J and K	.50	.50
All combined	.05	(not corrected)

N.B. All differences reaching the .02 level of confidence
have been underlined.

Transcription checked ?
Yes.

(c) Conclusions drawn from Tables XV and XVI showing the significance of the Score Fluctuations for weeks A, C D and E pooled, but eliminating week B

It will be seen from the above tables that eliminating week B produces a much more logically consistent picture. If attention is directed first to Table XV showing the results of the assessments on the Raw Scores, it will be seen that the only significant fluctuations are to be found in block A, in the scores of the E.G.S. on the E.S.O. Moreover, 4 out of 5 of these fluctuations are significant at the .001 level. There can be no question but that some factor or factors are operating selectively on the E.G.S. to cause fluctuations in their hits on the E.S.O.

The Scores Corrected for Bunching give a similar picture, but here the evidence is less strong, only the A hits returning a .01 P value, and the A,B,C,D & E classification a .001 P value.

Unless flaws can be found in the experimental procedure which selectively affected the E.G.S. hits on E.S.O. only, we can safely conclude that the correctness of the Linkage Material supplied to the E.G.S. caused them to score hits on the E.S.O., and not the C.S.O., whereas the incorrect Linkage Material supplied to the C.G.S. caused no such statistically significant increases in the scoring rate either on E.S.O. or C.S.O. This is an important conclusion, but it immediately raises additional questions which it would be interesting to have answered, such as which types of Linkage Material were the most effective, and the nature of the influence exerted.

At the moment we are in the position of a doctor who has prescribed a remedy with many ingredients, and who sees his patient respond to the medicine, but has no idea which of the ingredients are responsible for the recovery. It would

consequently be of value to consider each type of linkage separately, and if possible determine its contribution to the significant fluctuations observed. This will be attempted in the next assessment.

(6) Significance of the differences in scoring rate observed between week A and each other week of the experiment.

(a) Statistical Method:

The familiar χ^2 test is again used here, but this time the data is cast in a 2 X 2 table. With such 2 X 2 tables a difficulty arises because the χ^2 tables against which we assess our results are a continuous distribution, whereas our observed frequencies from which we calculate the value with which to enter the χ^2 tables are discontinuous. The continuous distribution is in fact the limit towards which the discontinuous observed distribution tends to approximate as the sample is made larger. In calculations involving more than one degree of freedom (i.e. having more than 4 cells) irregularities produced by small numbers are avoided by specifying that the test shall not be applied when the theoretically expected value in any cell falls below 5. With 2 X 2 tables a correction for continuity can be applied to small observed frequencies which improves the agreement of a discontinuous sample to the continuous χ^2 distribution against which it is being assessed. This correction, known as the Yates Correction, can be applied either by making the observed values less extreme by $\frac{1}{2}$ a unit, or by reducing the difference between the cross-products in such 2 X 2 tables by half the total number of observations, which Fisher shows amounts to the same thing. (102 p. 95). He points out that this slightly overcorrects the exaggeration of significance due to using a table of a continuous distribution.

A difficulty arises because there is no agreement among statistical authorities as to when it becomes unnecessary to correct for continuity. Thus Cochran, quoted by

Siegel, (107 p. 110), recommends using χ^2 corrected for continuity even when the total number of cases is greater than 40, and that the standard χ^2 test should not be used if the smallest expected frequency is less than 5. In such a case he advises the Fisher Exact Probabilities Test. Most authorities would agree that it is absolutely necessary to apply Yates' Correction when the observed values are small, and not necessary when they are reasonably big, but the majority remain discretely quiet as to what shall be regarded as big or small. As some decision had to be made, it was decided to follow the policy advocated by Rhine and Pratt, (106 p. 179, footnote), "For a more accurate estimate of the probability in evaluations that come out near the criterion of significance, Yates' correction for discontinuity should be made in obtaining χ^2 for tables having the observed value of any cell between 5 and 10. This consists of adding .5 to the smallest cell value and adjusting the other three cells to keep the marginal totals unchanged."

As an example of the method let us review calculation C 172, dealing with the Raw A hits scored by E.G.S. on E.S.O. in weeks A and B. The subjects scored 9 hits in week A in 1741 trials, and 10 in week B in 1760 trials. This gives a total of 19 hits in 3501 trials. On the null hypothesis that no significant difference exists, we would expect the hits scored in each week to be in proportion to the number of trials made, namely $\frac{1741}{3501}$ or .49729 of the total hits in week A, and $\frac{1760}{3501}$ or .50271 of the total hits in week B. This gives theoretically expected frequencies of 9.45 hits in week A, and 9.55 hits in week B. Applying Yates' Correction we add .5 to the 9 hits observed in week A, giving 9.5, and subtract .5 from the 10 hits observed in week B, which also gives 9.5.

The difference between these observed values, and the theoretically expected values is next calculated, and this deviation is .05 in both cases.

From this point on, the calculation of the χ^2 proceeds in the usual way. In this case, the obtained χ^2 for 1 degree of freedom is .00052, giving P lying between .99 and .98. There is thus no evidence to reject the null hypothesis we are testing, and we conclude that there is no significant difference in scoring rate between the two weeks. In this case it will be convenient to set our level of significance at .02 (two tail) in order to reject the null hypothesis. With the previous assessments made it was not possible to take into account the direction of the fluctuations. Our hypothesis is, however, that linkage caused an increase in scoring rate, except in week B, where we expect a decrease. Hence some readers may prefer to apply the probabilities for a one-tail test, and halve the P values shown. All the P values in the summary tables are for a two tail test. The assessments have been made both on the Raw Scores, and the Scores Corrected for Bunching.

It may be wondered why the formula usually recommended as most laboursaving was not employed for these calculations. Experiment showed that with the hand-operated calculating machine available it was less cumbersome and quicker to use the form of calculation which has been employed. Both methods give the same results.

(b) Calculations C 172 to C 211:

Tables XVII and XVIII:

Linkage Effects: Each week compared with week A to determine
Significance of score changes for Experimental Group Subjects.

Calculation
Serial No.: C 172.

Source data
Serial No.: D 1 & 16.

For Raw A hits on Experimental Series Originals, applying Yates' correction for continuity.

F	Week A .49729	Week B .50271	Totals: 1	Week A .50566	Week C .49434	Totals: 1
T	1741.0000	1760.0000	3501.0000	1741.0000	1702.0000	3443.0000
EH	9.45	9.55		13.65	13.35	
OH	9.5	9.5	19.0	9.5	17.5	27.0
D	.05	.05		4.15	4.15	
D ²	.0025	.0025		17.2225	17.2225	
$\frac{D^2}{EH}$.00026	.00026		1.26	1.29	
EM	1731.55	1750.45		1727.35	1688.65	
$\frac{D^2}{EM}$.0000	.0000		.01	.01	
χ^2 (for 1 d.f.) =			.00052			2.57
P two tail lies between:	.99 and .98			P two tail lies between:		.20 and .10
F	Week A .50566	Week D .49434	Totals: 1	Week A .50921	Week E .49079	Totals 1
T	1741.0000	1702.0000	3443.0000	1741.0000	1678.0000	3419.0000
EH	10.62	10.38		15.28	14.72	
OH	9.5	11.5	21.0	9.5	20.5	30.0
D	1.12	1.12		5.78	5.78	
D ²	1.2544	1.2544		33.4084	33.4084	
$\frac{D^2}{EH}$.1181	.1208		2.19	2.27	
EM	1730.38	1691.62		1725.72	1663.28	
$\frac{D^2}{EM}$.0007	.0007		.02	.02	
χ^2 (for 1 d.f.) =			.2403			4.50
P two tail lies between:	.70 and .50			P two tail lies between:		.05 and .02

N.B.: + indicates an improvement on week A scores,
- indicates a decrease from week A scores. (See table XVII)

Calculation checked ? Yes.

TABLE XVII

SIGNIFICANCE OF THE RAW SCORING RATE DIFFERENCES
BETWEEN WEEK A AND EACH OTHER WEEK SEPARATELY.

(Calculations C 172 to C 199).

- N.B. (1) An improvement on week A scoring rates is shown by a +
A decrease from week A scoring rate is shown by a -
- (2) All significances shown are for P two tail. P is always smaller than the value shown.
- (3) P values reaching .02 two tail have been underlined.
- (4) In cases marked * the theoretically expected frequencies were too small for accurate calculation.

Experimental Group Subjects' hits on								
Hit type:	(A) E.S.O.				(B) C.S.O.			
	Week A and:				Week A and:			
	B	C	D	E	B	C	D	E
A	+ .99	+ .20	+ .70	+ .05	- *	+ *	+ *	+ *
A, B, C, D & E	- .80	+ .70	+ .001	+ .001	- .001	+ .70	- .50	- .30
All F, G + H3	- .10	+ .02	+ .20	+ .02	- .05	+ .70	+ .70	+ .20
All F, G & H	- .95	+ .02	+ .05	+ .01	- .50	+ .80	+ .80	+ .80
I & K	- .30	- .70	+ .001	+ .05	- .05	+ .50	+ .70	+ .70
I, J & K	+ .30	- .70	+ .001	+ .02	- .001	+ .50	+ .80	+ .80
All combined	- .90	+ .30	+ .001	+ .001	- .001	+ .80	- .90	- .90

Control Group Subjects' hits on								
Hit type:	(C) E.S.O.				(D) C.S.O.			
	Week A and:				Week A and:			
	B	C	D	E	B	C	D	E
A	- .50	+ .90	- .70	+ .98	- *	- *	+ *	- *
A, B, C, D & E	- .50	+ .20	+ .20	+ .30	- .05	- .70	- .90	- .30
All F, G + H3	- .01	- .95	+ .50	+ .30	- .20	+ .70	+ .80	+ .70
All F, G & H	- .20	- .98	+ .20	+ .10	+ .90	+ .10	+ .05	+ .10
I & K	+ .10	+ .30	- .90	+ .20	+ .70	+ .50	+ .95	+ .70
I, J & K	+ .30	+ .20	+ .90	+ .50	- .50	+ .10	+ .20	+ .50
All combined	- .50	+ .20	+ .10	+ .01	+ .95	+ .20	+ .10	+ .20

Transcription checked ? Yes.

TABLE XVIII

COMPARISON OF SIGNIFICANCES OBTAINED FROM RAW SCORES & SCORES
CORRECTED FOR BUNCHING ON EXPERIMENTAL SERIES ORIGINALS ONLY.
DIFFERENCES BETWEEN WEEK A AND EACH OTHER WEEK.

(Calculations C 172 to C 185 and C 200 to C 211)

- N.B. (1) An improvement on week A scoring rates is shown by a +
A decrease from week A scoring rates is shown by a -
- (2) All significances are for P two tail. P is always
smaller than the value shown.
- (3) P values reaching .02 two tail have been underlined.

Experimental Group Subjects' hits on E.S.O.								
Hit type:	(A) Raw Scores:				(B) Corrected Scores:			
	B	Week A and			B	Week A and		
		C	D	E		C	D	E
A	+ .99	+ .20	+ .70	+ .05	+ .50	+ .02	+ .20	+ <u>.01</u>
A, B, C, D & E	- .80	+ .70	+ <u>.001</u>	+ <u>.001</u>	- .20	+ .50	+ <u>.02</u>	+ <u>.001</u>
All F, G + H3	- .10	+ <u>.02</u>	+ .20	+ <u>.02</u>	- .10	+ <u>.02</u>	+ .50	+ .05
All F, G & H	- .95	+ <u>.02</u>	+ .05	+ <u>.01</u>	- .20	+ <u>.02</u>	+ .20	+ <u>.01</u>
I & K	- .30	- .70	+ <u>.001</u>	+ .05	- <u>.01</u>	- .20	- .95	- .95
I, J & K	+ .30	- .70	+ <u>.001</u>	+ <u>.02</u>	- <u>.01</u>	- .30	- .90	+ .90
All combined	- .90	+ .30	+ <u>.001</u>	+ <u>.001</u>	(not corrected).			

Control Group Subjects' hits on E.S.O.								
Hit type:	(C) Raw Scores				(D) Corrected Scores			
	B	Week A and			B	Week A and		
		C	D	E		C	D	E
A	- .50	+ .90	- .70	+ .98	- .98	+ .50	+ .99	+ .70
A, B, C, D & E	- .50	+ .20	+ .20	+ .30	- .50	+ .10	+ .70	+ .50
All F, G + H3	- <u>.01</u>	- .95	+ .50	+ .30	- <u>.01</u>	- .95	+ .98	+ .30
All F, G & H	- .20	- .98	+ .20	+ .10	- .05	- .98	+ .70	+ .10
I & K	+ .10	+ .30	- .90	+ .20	- .98	+ .70	- .20	+ .80
I, J & K	+ .30	+ .20	+ .90	+ .50	- .70	+ .50	- .50	- .90
All combined	- .50	+ .20	+ .10	+ <u>.01</u>	(not corrected).			

Transcription checked ?
Yes.

(c) Conclusions drawn from Tables XVII and XVIII on the significance of scoring rate differences between week A and each other week.

Let us consider first the Raw Score results given in Table XVII. In this and subsequent tables a plus before a P value indicates that an improvement on the scoring rate established in week A took place, whereas a minus indicates that a falling off in scoring rate from that shown in week A took place. This information has been obtained from the data sheets giving the expected scores per 2,000 trials per week per subject group.

Viewing the table as a whole, it will be seen that as with the previous assessment, the most consistent and significant fluctuations have taken place in block A, that is, in the hits scored by E.G.S. on E.S.O. Moreover, all these significant changes are positive, showing an increase in the scoring rate when linkage was given. In block B the E.G.S. also show three significant decreases in hits scored on the C.S.O. in week B, when they were supplied with squares of colour as linkage material.

By contrast the C.G.S. show no significant changes at all in their hits on the C.S.O. in block D, and only two significant fluctuations in their hits on E.S.O. in block C, a decrease in week B for the "All F, all G & H3 hits" and an increase in week E in the final "All combined" category.

Table XVIII shows that correcting for bunching slightly changes the picture, two new significant decreases appearing in week B in block B, for hits by E.G.S. on E.S.O.

Returning once more to Table XVII, it would appear that a well-marked difference exists between block A and the other three blocks of data. Let us now examine these results week by week in terms of the linkage material provided:

Week B: The Colour Square Linkage: During this week the agent was given squares of colour on his sheets for the target drawings, and was asked to visualize each drawing he made as if drawn in the colour of the square provided. The E.G.S. were given identical squares of colour, the C.G.S. incorrect squares of colour. Subjects from both groups complained that the colours hindered their getting impressions. Now turning to the data of Table XVII, we find, quite unexpectedly, that no significant changes have taken place except a fairly consistent and highly significant decrease in the number of hits scored by the E.G.S. on C.S.O. in block B. (There is one further significant decrease in block C for the "All F, all G & H3" category, but it is quite atypical and can be ignored.) Now these negative deviations in block B cannot be due merely to the constraining effect of the squares of colour, because this effect would then also show up in block D. But this has not occurred. It would seem that they must be attributed to the correctness factor of the colour squares supplied to the E.G.S. The influence of the colours would thus appear to have been to steer the E.G.S. away from the control targets not bearing any colours at all, but not to aid them particularly in scoring hits on the E.S.O. There is no evidence at all that the correctness factor or the suggestion factor postulated in the colour linkage aided the subjects in scoring hits. This finding, taken in conjunction with the findings of the Second Main Analysis, does not support the association-to-association type of linkage mechanism posited by Carington's Association Theory of Telepathy.

Week C: The Psychometric Handkerchief Link:

Only two significant deviations occur in week C, in block A dealing with the E.G.S. hits on E.S.O. They both only reach the .02 (two tail) level, but as they are both positive,

they might be considered as reaching the .01 level. If they are accepted at their face value they would suggest that the provision of psychometric handkerchief links did help the subjects to make contact with the desired agent. Since the C.G.S. do not show any similar deviations, the effectiveness would appear to be due, at least in part, to the fact that the handkerchiefs had actually been used by the agent, rather than solely to any possible suggestion effects that might have been operating.

Week D: Symbolic Descriptive Linkage: Photograph of Agent and Personality Material on him:

Here we are on much firmer ground, having no less than four positive deviations reaching the .001 level of significance in the E.G.S. hits on E.S.O., and no other significant fluctuations in any of the other blocks of data. This suggests that the provision of the photograph of the agent, and the personality material correctly describing him was effective in increasing the number of hits scored; and that this effectiveness was due, again at least in part, to the correctness factor.

Week E: Symbolic-descriptive Linkage, with Agent only using Squares of Colour:

In week E all but one of the significant increases are confined to E.G.S. hits on E.S.O. depicted in block A. (The one exception occurs in block C, showing C.G.S. hits on E.S.O. for the "All combined" category, and is so atypical of the rest of block C that it is proposed to ignore it as a chance produced effect, such as can be expected in terms of probability theory in a table presenting 104 calculations of significance. Alternatively; it may be caused by the normal spurt in the number of hits at the end of an experiment, although if so, one would expect the other hit categories to show it too.) It will be seen that all the deviations in this final week are positive in block A,

and that five out of the seven presented are sufficiently marked to be regarded as statistically significant. Here too the crucial factor would appear to be the correctness of the linkage material supplied, over and above any possible benefit accruing from its suggestion value.

Before ending this review of this section of the data it would be as well to return to Table XVIII to check on the changes produced when the Raw Scores are corrected for bunching errors. The only major difference would seem to be the appearance of two negative deviations significant at the .01 level (two tail) in the E.G.S. hits on E.S.O. in block B in week B. In other respects the trends are similar to those shown in the Raw Scores, but less clear-cut.

Let us next consider the evidence bearing on the existence of the suggestion factor and the correctness factor of the linkage. The results obtained clearly demonstrate the existence of an effective correctness factor operating to increase the number of hits scored. By contrast our analyses provide no clear evidence of the existence of a suggestion factor capable of significantly influencing the results. If extant such a factor should have produced significant changes in the scoring rates of the C.G.S. on the E.S.O., but no such consistent changes are observable. It may be tempting to infer the existence of a suggestion factor on common-sense grounds, and an inspection of the graphically presented scoring rates may lend some little colour to the supposition, but its existence in this experiment has not been demonstrated with any statistical certainty.

An interesting feature of the results is that the significant changes have occurred in hit categories other than A. Taken at its face value this would suggest that the linkage exerts its effect not by increasing the "direct hits"

scored, but by increasing the number of partial or distorted hits. However this conclusion is perhaps not warranted. The effect may be due to the relatively small numbers of A hits. An experiment on a larger scale might reveal significant increases in the A hits too, but the significances actually obtained hold out little hope of this.

So far in these assessments we have examined the scores returned by each subject group on each series of originals separately, to find whether any show fluctuations which are statistically significant. We found such extra-chance fluctuations only in the case of the E.G.S., and we discussed possible explanations of these in terms of the suggestion factor and the correctness factor operating in the linkage material effects. Reflection will show, however, that much of the argument used rests on the implicit assumption that the two subject groups were initially equal in hit scoring ability. The size of the groups and the way they were randomly selected would seem to make this assumption fairly plausible, but for safety it would be better to test it by comparing the scoring rates returned by the two subject groups in week A to find whether they differ significantly or not.

The following example will perhaps make the crucial nature of the assumption clearer. Suppose that the two subject groups were initially not equal in hit-scoring ability, and to take an extreme case for clearer illustration, suppose too that all the good E.S.P. subjects were allocated to the Experimental Group, whereas the Control Group contained no subjects capable of exercising E.S.P. The incorrect linkage issued to the C.G.S. would then have no effect on their scoring rates, because of their lack of E.S.P. ability, and not because the suggestion factor was ineffective. In turn the significant fluctuations found in the case of the E.G.S. could not be attributed merely to the addition of the correctness factor to the linkage. It could still be due in part or in whole to

the influence of suggestion. Because of the inadequacy of the control group we would not be able to distinguish between it and the correctness factor. It becomes necessary therefore, to assess the significance of the scoring rate differences between the two subject groups.

(7) Comparison of the Scoring Rates of the E.G.S. with those of the C.G.S. on the E.S.O. only, to find whether they show significant differences.

(a) Introduction:

We have already discussed the necessity of establishing whether or not the two subject groups were initially equal in hit scoring ability in week A. If no significant differences are found to exist in week A, but significant differences crop up in subsequent weeks in favour of the E.G.S., this would shed additional light on the relative importance of the suggestion and correctness factors. Such results would suggest that the correctness factor is more important than the complex of suggestion factor and all other uncontrolled factors as well, since correctness by itself is capable of producing a significant increase, whereas the suggestion factor complex failed to produce such significant changes in the C.G.S. hits on E.S.O.

(b) Statistical Method:

This is similar to that employed in calculations C 172 to C 211, except that instead of comparing week A scores with week B scores for a particular subject group, we compare E.G.S. hits in week A with C.G.S. hits in week A. This time too, the conventional formula has been employed, as used previously in calculations C 88 to C 103, in connection with the Psychic Questionnaire material. Yates' Correction has been applied in cases where the observed values fall between 10 and 5

(c) Calculations C 212 to C 218:

Table XIX:

Significance of Weekly Scoring Rate Differences between Control and Experimental Group Subjects.

Calculation Serial No.: C 212. Source Data Serial No.: D 1.
 For Raw A hits on Experimental Series Originals.

Week A	EGS		CGS		Totals:		
Hits	A	9	B	14	z	23	$m = AD-BC = 8,165$
Misses	C	1,732	D	1,787	y	3,519	$\frac{1}{2}N = 1,771$
Totals	w	1,741	x	1,801	N	3,542	$m - \frac{1}{2}N = 6,394$

$$\frac{m^2 N}{wxyz} = \frac{(m - \frac{1}{2}N)^2 N}{wxyz} = \frac{144,808,421,912}{253,781,281,917}$$

$$\chi^2 = .5706 \quad P = .50 \text{ \& } .30$$

Week B	EGS		CGS		Totals:		
Hits	A	10	B	9	z	19	$m = AD-BC = 1,760$
Misses	C	1,750	D	1,751	y	3,501	$\frac{1}{2}N = 1,760$
Totals	w	1,760	x	1,760	N	3,520	$m - \frac{1}{2}N = 0$

$$\frac{m^2 N}{wxyz} = \frac{(m - \frac{1}{2}N)^2 N}{wxyz} = \frac{0}{206,049,254,400}$$

$$\chi^2 = \quad P = 1.00 \text{ \& } .99$$

Week C	EGS		CGS		Totals:		
Hits	A	18	B	15	z	33	$m = AD-BC = 6,132$
Misses	C	1,684	D	1,744	y	3,428	$\frac{1}{2}N = 1,730.5$
Totals	w	1,702	x	1,759	N	3,461	$m - \frac{1}{2}N$

$$\frac{m^2 N}{wxyz} = \frac{130,138,528,464}{338,672,667,432}$$

$$\frac{(m - \frac{1}{2}N)^2 N}{wxyz} = \quad$$

$$\chi^2 = .3843 \quad P = .70 \text{ \& } .50$$

Week D	EGS		CGS		Totals:		
Hits	A	12	B	11	z	23	$m = AD-BC = 2,482$
Misses	C	1,690	D	1,756	y	3,446	$\frac{1}{2}N = 1,734.5$
Totals	w	1,702	x	1,767	N	3,469	$m - \frac{1}{2}N$

$$\frac{m^2 N}{wxyz} = \frac{21,370,163,956}{238,363,203,972}$$

$$\frac{(m - \frac{1}{2}N)^2 N}{wxyz} = \quad$$

$$\chi^2 = .0897 \quad P = .80 \text{ \& } .70$$

Week E	EGS		CGS		Totals:		
Hits	A	21	B	14	z	35	$m = AD-BC = 13,678$
Misses	C	1,657	D	1,756	y	3,413	$\frac{1}{2}N = 1,724$
Totals	w	1,678	x	1,770	N	3,448	$m - \frac{1}{2}N = 11,954$

$$\frac{m^2 N}{wxyz} = \frac{645,078,334,432}{354,788,517,300}$$

$$\frac{(m - \frac{1}{2}N)^2 N}{wxyz} = \frac{492,753,921,360}{354,788,517,300}$$

$$\chi^2 = 1.818 \quad P = .20 \text{ \& } .10 \quad \chi^2 = 1.39 \quad P = .30 \text{ \& } .20$$

Calculation checked ? Yes.

TABLE XIX

SIGNIFICANCE OF WEEKLY SCORING RATE DIFFERENCES
BETWEEN CONTROL AND EXPERIMENTAL GROUP SUBJECTS
RAW HITS ON EXPERIMENTAL SERIES ORIGINALS.

(Calculations C 212 to C 218)

- N.B. (1) A + indicates that the E.G.S. returned the better score
 A - indicates that the C.G.S. returned the better score
- (2) An * indicates cases where a significant increase in
 E.G.S. hits on E.S.O. was found in previous assessments
- (3) Significances reaching the .02 level (two tail) have
 been underlined.
- (4) P is always smaller than the value shown.

<u>Hit Type</u>	<u>Week A</u>	<u>Week B</u>	<u>Week C</u>	<u>Week D</u>	<u>Week E</u>
A	-.50	-1.00	-.70	+.80	+.20
A,B,C,D & E	-.70	-.80	+.20	-.20*	<u>+.01*</u>
All F, G +H3	-.20	-.80	+.20*	-.70	-.98*
All F, G & H	-.20	-.95	+.30*	-.70	-.90*
I & K	+.80	<u>-.01</u>	-.20	<u>+.001*</u>	+.30
I, J & K	+.90	+.95	-.10	<u>+.001*</u>	+.10*
All combined	-.95	+.70	-.70	<u>+.02*</u>	+.20*

Transcription checked ?
 Yes.

(d) Conclusions drawn from Table XIX: Significance of Weekly Raw Scoring Rate Differences between C & E.G.S. on Experimental Series Originals.

If attention be directed to week A, it will be seen that none of the differences there given reach the .02 (two tail) level of significance. We can conclude that the two subject groups were thus not significantly different in hit scoring ability under conditions of no linkage in week A. This establishes the crucial point on which the argument already outlined hinges. We can therefore legitimately compare the scores of the two groups in subsequent weeks.

With regard to the evidence bearing on the relative importance of the suggestion and correctness factors, the results tend to be suggestive rather than conclusive. Cases where the E.G.S. previously showed significant increases on their week A scoring rates have been marked with an asterisk. Only in the case of weeks D and E is there any suggestion that the correctness factor by itself is sufficiently strong to produce significant increases in scoring rate. Even here, the evidence is weakened by the lack of consistency from week D to week E. As the linkage material used in these two weeks was almost identical, one would expect a significant difference shown in week D to carry over into week E as well, but this has not occurred.

Significant differences in the expected direction have failed to occur in weeks B and C, employing colour square linkage and psychometric handkerchief links. This would seem to indicate that in these types of linkage the correctness factor was less powerful than in the symbolic-descriptive photograph and personality description linkage.

It seems that we have pushed our analysis of the data pretty well as far as is justified by the design of the experiment, and the nature of the material gathered.

It is perhaps inevitable that an investigation such as this raises more questions than it answers. It would be instructive to know, for instance, the exact relative importance of the various factors operating in the linkage material, and whether the suggestion factor does exist. It would be useful to know the role played by the decline effect in an experiment of this kind. One would further like to know whether the linkage continues to exert its effect even if it is removed after contact has been established between agent and subjects, and whether it operates as an all-or-none phenomena, or whether a cumulative improvement can be obtained by adding more and more linkage from week to week. Unfortunately the design of this experiment is too crude to be able to answer these questions with any certainty. New experiments are needed specifically designed to investigate these points, perhaps employing a randomised block technique, or one of its modifications. Despite its limitations this experiment has demonstrated that the provision of certain types of linkage material effectively increases the subjects' ability to score hits, and has resulted in other interesting findings which are summarised below.

(8) Summary of the Conclusions drawn from the Fourth Main Analysis of the Linkage Data:

(a) It was found that no significant difference in hit scoring ability existed between the Control and Experimental Group of subjects before the provision of linkage material.

(b) When the subjects were provided with a photograph of their agent, and a description of his personality consistent significant improvements in their scoring rates took place. Provision of an equated but incorrect set of linkage material to the Control Group Subjects caused no such significant improvement.

(c) Provision of a psychometric handkerchief link produced two significant increases, but these were neither so consistent nor so significant as those produced by the photograph and descriptive material of the agent. Incorrect handkerchief links supplied to the C.G.S. produced no such significant increases.

(d) Provision of associative links in the form of squares of colour caused no significant increases in scores for either subject group, but did produce a highly significant and fairly consistent decrease in the hits scored by the E.G.S. on the C.S.O. By contrast the C.G.S., who received incorrect squares of colour, showed no such decrease in hits scored on the C.S.C.

It is suggested that the correct colours acted by guiding the E.G.S. away from originals totally unassociated with any squares of colour, and not by linking them directly with the target originals bearing the correct colours

(e) The Second Main Analysis showed that there was no tendency for more hits to be scored on originals during the weeks in which they were actually being used as targets than in other weeks, but that the hits on any original were scattered evenly throughout the whole experimental period. Taken in conjunction with the evidence from this Fourth Main Analysis, all the indications are against the hypothesis that the linkage material acted by guiding the subjects to specific items in the agent's mind, as postulated by Whately Carington in his Association Theory of Telepathy. It would seem rather that it exerted its effect by giving the subjects access to the agent's mind as a whole. Even week B, using colour links, showed no trace of item-to-item effect. Once linkage had been established, the whole of the agent's mind, past, present and future was apparently equally available to the subject, confirming Hettinger's observations.

(f) By contrast, attention is drawn to the fact that at the subjects' end of the linkage process, the ordinary laws of before and after causality applied. Before linkage was given, scores were low: after linkage was provided a significant improvement in scoring rate took place. Linkage thus did not help the subjects before it was given. If its effectiveness had shown any tendency to spread in time, no significant increases in scoring rate would have been found in the experiment. To reveal such an effect two separate but equated experiments would be necessary, using different subjects, targets and agents in each. Comparison of the results of two such experiments would show significantly better scoring rates in the one in which linkage material had been provided, than in the other without linkage.

(g) The analysis has shown that in the case of the photograph and personal description linkage, its effectiveness depends at least in part on its correctness. Other components such as the suggestion value may or may not be contributing to this effectiveness, but alone all such other factors combined did not produce any significant increases in score comparable to those produced when the correctness factor was added.

(h) It has been pointed out that further problems concerning the linkage process have been raised by this experiment, *inter alia*: the number and relative importance of various factors such as the correctness of the linkage, its suggestion value, the decline effect, and so on; whether the effect exerted by the linkage is cumulative or not, and many others. It is suggested that specially designed experiments will be needed to obtain answers to these questions.

This completes the statistical assessment of the data, and the next section will deal with a qualitative assessment of a sample of the results returned by the subjects.

FIFTH MAIN ANALYSISQUALITATIVE EVALUATION OF A SAMPLE
OF DRAWINGS RETURNED BY THE SUBJECTS.

(1) <u>Introduction:</u>	327
(2) <u>How to find when a drawing was made, or a target used:</u>	329
(3) <u>Summary of the Assessment Code</u>	329
(4) <u>List of Control and Experimental Series Originals:</u>	330
(5) <u>Reproductions of hits on Experimental Series Originals:</u>	332
(6) <u>Reproductions of hits on Control Series Originals:</u>	362
(7) <u>Classification of subjects' drawings into idea and shape resemblances:</u>	376
Table XX: E.S.O. data.	376
Table XXI: C.S.O. data.	378
(8) <u>Conclusions drawn from Tables XX & XXI, and from an examination of the data:</u>	379

(1) Introduction:

In this section a sample of the drawings returned by the subjects is presented, together with the judges' assessments of them. It is hoped that this material will illustrate for readers both the type of data collected, and the method of awarding hits; and that it will to some extent complement the more abstract assessments of statistical significance which have been our concern up to now. Since these statistical tests have shown that some of the hits scored can validly be attributed to the operation of E.S.P. we can now legitimately undertake a qualitative study of the data in the hope that such a study will reveal additional evidence bearing on the nature of the E.S.P. process which occurred. It is not intended to undertake a full-scale investigation of this kind in this report, but merely to draw attention to interesting aspects of the data. A complete analysis of the material may possibly form the subject matter of a subsequent report. here we shall limit ourselves mainly to a consideration of whether the drawings appear to have been made with the aid of telepathy or of clairvoyance, and to observations on the kinds of distortions which they show.

To obtain information on the first of these problems, the subjects' drawings have been classified by inspection into four categories:

- (a) Cases where the subjects have reproduced the idea correctly, but the shape incorrectly.
- (b) Cases where the shape has been well reproduced, but the idea mistaken.
- (c) Cases with correct idea and good shape.
- (d) Doubtful cases.

It is suggested that the reader examine the drawings which have been reproduced in the following pages, and attempt to

decide for himself whether the subjects appear to have been guided by the apprehension of symbols and ideas in making their drawings, or by what appears to be direct visual apprehension of the targets.

In selecting the sample, all the A title hits have been included, as well as the strict shape hits in the F3 category. Some examples of each of the other types of hit have been included as well. The target original is always enclosed in a rectangle or square to mark it off from the drawings made by the subjects. Each drawing made by a subject bears a serial number, which can be used to find when the drawing was made by referring to the information given on the next page. This serial number is followed by the code assessments of the three judges, which can be interpreted by reference to the summary which will also be found overleaf. The reproductions of the drawings are preceded by a list showing the titles of originals in the Experimental and Control Series, which can be used to find when any particular original was used as a target.

(2) How to find when a drawing was made, or a target used:

Each subject's drawing is accompanied by a serial number, which can be used to find which subject group its originator belonged to, and in which section of the experiment it was made, by referring to the table below:

<u>Week:</u>	<u>Experimental Group Subjects'</u> <u>Serial Numbers:</u>	<u>Control Group Subjects'</u> <u>Serial Numbers:</u>
A	1 to 1820	1821 to 3999
B	4 and 5 thousand	6 and 7 thousand
C	8 and 9 thousand	10 and 11 thousand
D	12 and 13 thousand	14 and 15 thousand
E	16 and 17 thousand	18 and 19 thousand.

In addition, drawings whose serial numbers end in 1 and 2 were probably made on Monday, those ending in 3 and 4 on Tuesday, 5 and 6 on Wednesday, 7 and 8 on Thursday, and 9 and 10 on Friday.

A list showing the correct titles of targets in both C and E groups is given overleaf. The E.S.O. have been arranged in the chronological order in which they were used as targets, the first pair in any week having been made on Monday, the second pair on Tuesday, and so on.

(3) Summary of the Assessment Code:NAME:

- A = Name wholly right: ("Fountain" for "Fountain")
- B = Name generalized: ("Mug" for "Beer mug")
- C = Name specialized: ("Beer mug" for "Mug")
- D = General name right, but different species:
("Beer mug" for "Tooth mug")
- E = Synonym: ("Car" for "Automobile")

CONTENT OF DRAWING:

- F = Original reproduced without additions or omissions..
- G = Part given, whole reproduced.
- H = Whole given, part reproduced.

Add if distorted: *
Add if elaborated: o

Indicate the degree of resemblance thus: 3 = Good
2 = Fair
1 = Faint

IDEA OR ASSOCIATION:

- I = Clear association, (though name and shape different)
(E.g.: "Wind" for "Weather vane")
- J = Idea right, name wrong. (E.g.: "Chrysalis of Man"
for "Mummy")
- K = Miscellaneous resemblances not falling under any of the
the above.

Assessments are recorded as follows:

Serial Number of Per- cipient's drawing:	Number of Original it resembles:	Code analysis of resemblance
---	-------------------------------------	---------------------------------

17790

36

A F * 3

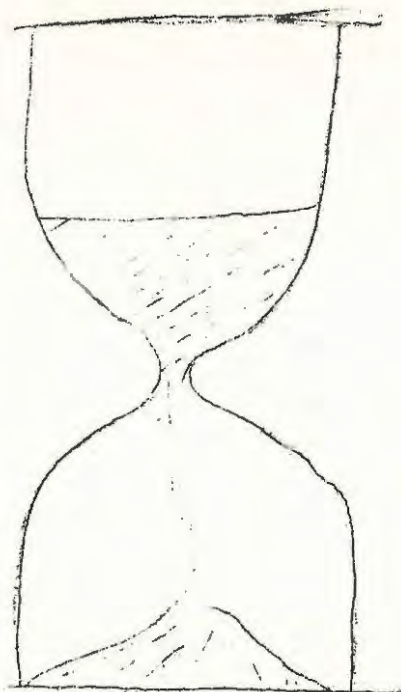
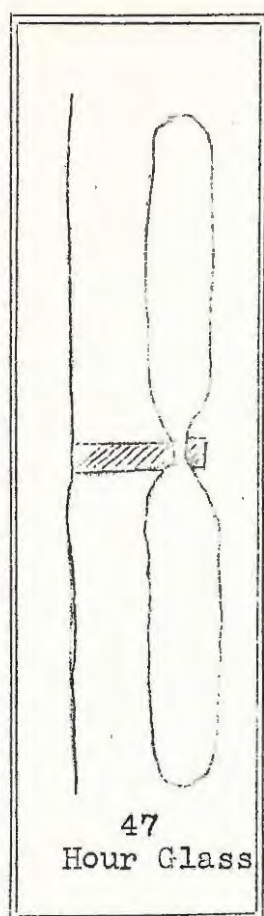
(4) LIST OF CONTROL AND EXPERIMENTAL
SERIES ORIGINALS.

- Notes: (1) An asterisk next to a title indicates that it occurred twice in the 15,835 suggestions returned in the Preliminary Frequency Experiment.
- (2) Experimental Series Originals are given in the order in which they were drawn by the agent during the course of the experiment.
- (3) Control Series Originals are listed in alphabetical order.
- (4) The double hyphen is intentional.

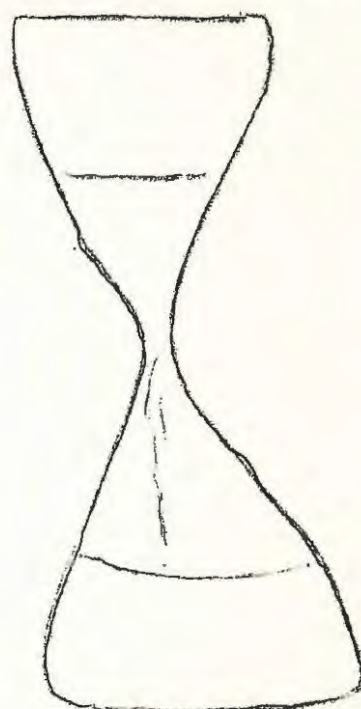
	<u>EXPERIMENTAL SERIES ORIGINALS</u>	<u>CONTROL SERIES ORIGINALS.</u>
Section A:	BOTTLE OPENER	ACCORDEON
	ANVIL	ALADDIN'S LAMP
	BUS STOP	ANT BEAR
	FIRE EXTINGUISHER	ARC DE TRIOMPHE
	VICE *	BAG=PIPES
	SNAIL	BILLIARD TABLE
	ACE OF HEARTS	BELLOWS
	HOUR GLASS	CAR JACK
	SAINT	CRUTCHES
	MORSE KEY	CUPID
Section B:	FLIT SPRAY	DECANTER
	TURNSTILE	DISCUS THROWER (MODEL)
	FLAT IRON	DUST PAN
	HOE	EARPHONES
	DUMB=BELL *	ECC=BEATER
	WISH=BONE	EGYPTIAN MUMMY
	SPIFFY	EIFFEL TOWER
	JAVELIN	FATHER CHRISTMAS
	CANDLE STICK	FISH MOTH
	BEACON *	FISHING GAIT
Section C:	PENGUIN *	FRENCH CURVE
	T=SQUARE	GOD PAN
	VOLCANO	GOLLYWOG
	SEE=SAW	GAS MASK

EXPERIMENTAL SERIES ORIGINALSCONTROL SERIES ORIGINALSSection C
(continued)PENDULUM
ROLLING-PIN
TRIDENT
TIN OPENER
ASTERISK
FEZHARROW
HASP AND STAPLE
HEDGEHOG
JODPHURS
KILT
LAUREL CROWNSection
D:BARBED WIRE BARD *
U-TUBE
SENTRY BOX
WING-NUT
ROSETTE
SEMI-QUAVER
KNOBKERRIE
MR. CHAD
WHALE
BISHOP'S MITRELEOPARD
LORGNETTE
ORGAN *
PEACOCK
PHOENIX
PRICKLY PEAR
RICKSHA
PUNCH BAG
RADAR AERIAL
PRIMUS STOVESection
E:(ARTIST'S) PALETTE
VIOLIN BOW
BOOMERANG
HONEYCOMB
CRICKET PADS
PETROL PUMP *
TAJ MAHAL
SATURN
RHINOCEROS
WEATHERCOCKROCKING HORSE
ROLLER SKATES
SCORPION
SCOUT BADGE
SHOE-TREE
SEAL
STAPLING MACHINE
SPHINX
W.C.
WATERBAG .

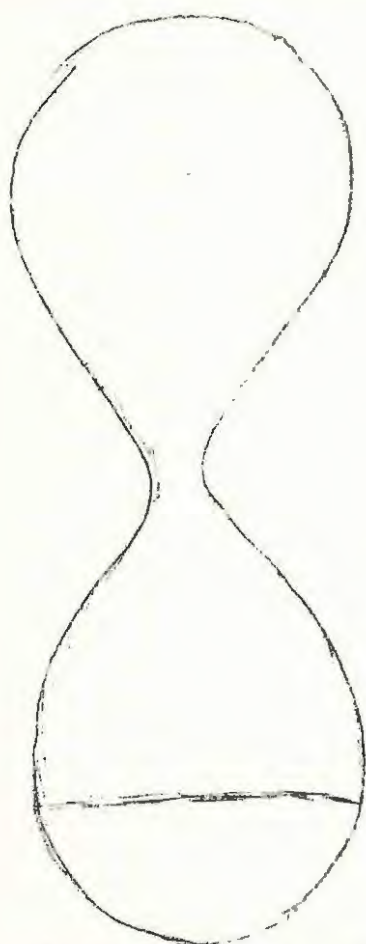
(5) HITS ON EXPERIMENTAL SERIES
ORIGINALS



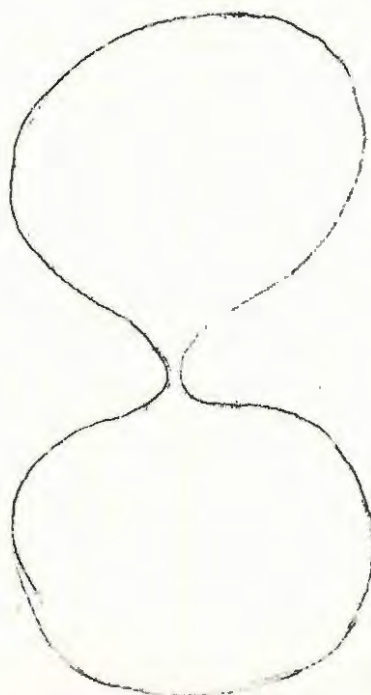
16084 "Hour glass"
AHox1, AH3, AH2.



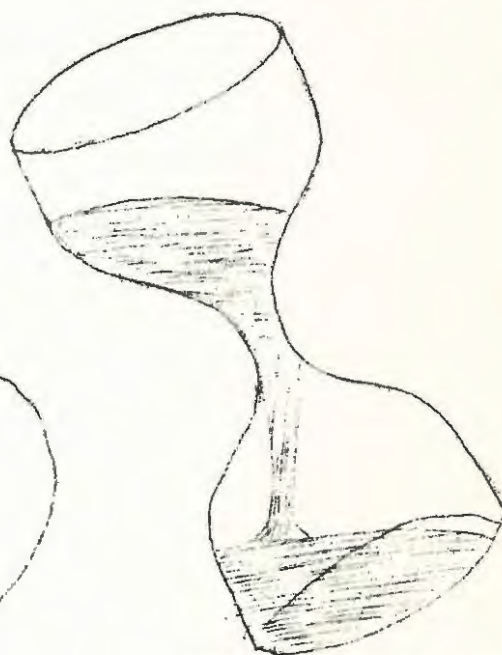
16566 "Hour glass"
AHox1, AH3, AH2.



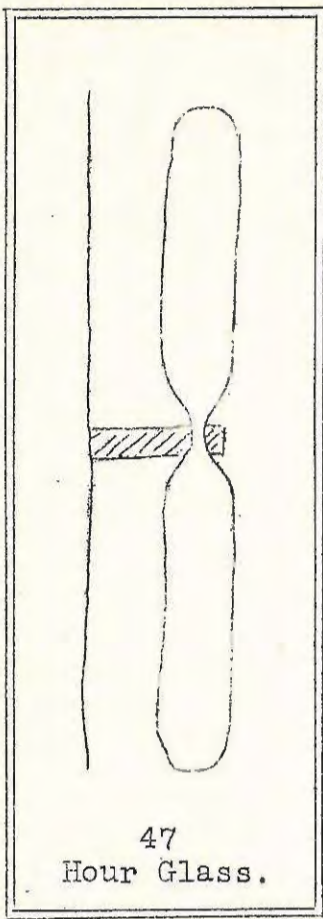
12991 "Hour-Glass"
AHox1, AH3, AH2.



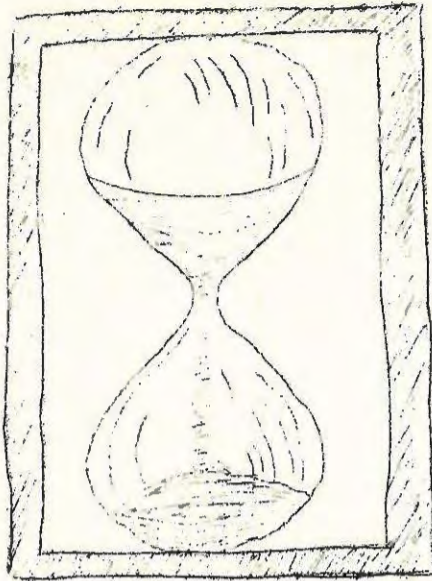
19542 "Hour glass"
AHx1, AH3, AH1.



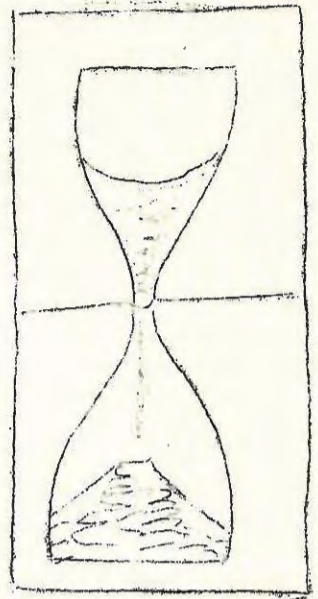
9822 "Hour Glass"
AHo1, AH3, AH2.



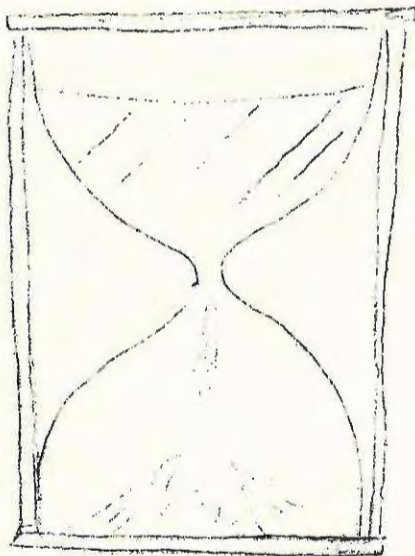
47
Hour Glass.



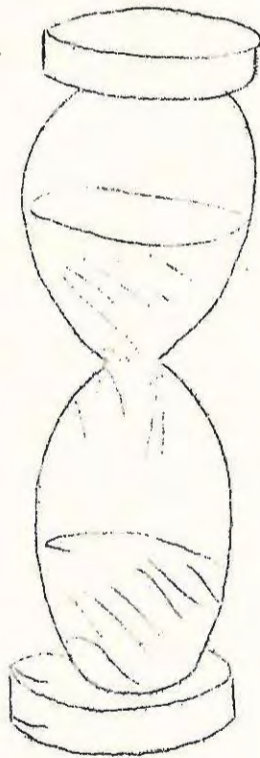
10653 "Hour Glass"
AFo2, AH3, AH2.



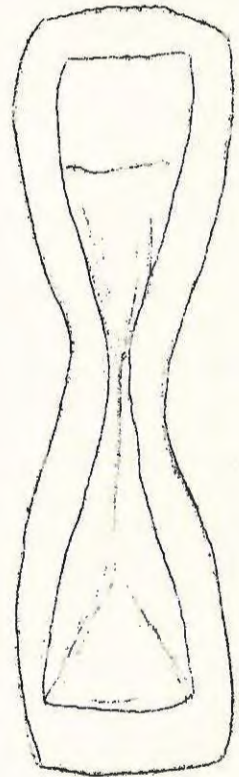
2874 "Hour Glass"
AFox3, AH3, AH2.



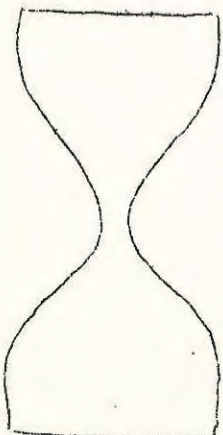
18154 "HourGlass"
AHox2, AH3, AH2.



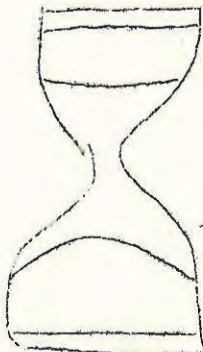
5114 "Hourglass"
AHox1, AH3, AH2.



7643
"Egg Timer"
EHx3, AH3, EH2.



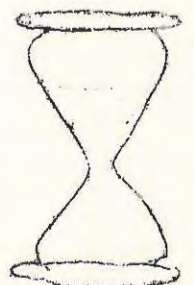
10581 "Hour-Glass"
AHx1, AH3, AH2.



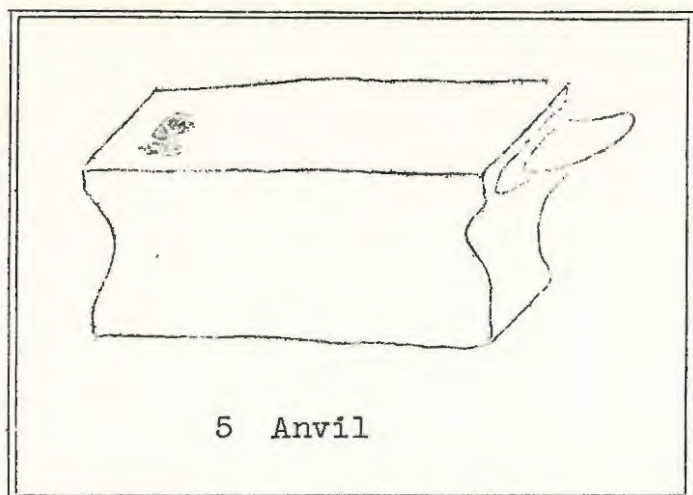
2061 "Hourglass"
AHx1, AH3, AH2.



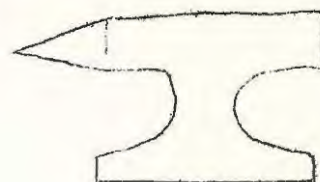
19788
"Possibly an Hourglass"
AHox1, AH3, AH2.



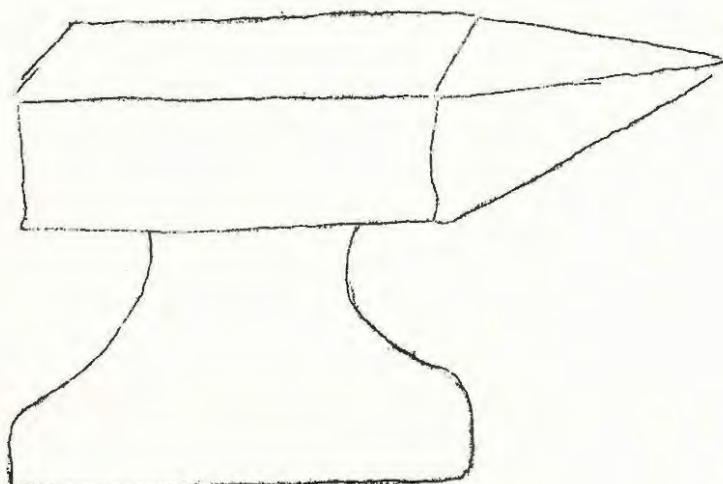
8070
"Hour-glass"
AFox2, AFx2, AFx3



5 Anvil



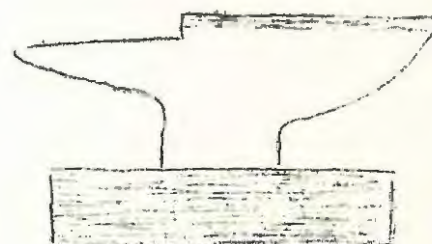
6597
"Anvil"
AFx3, AF2,
AF2.



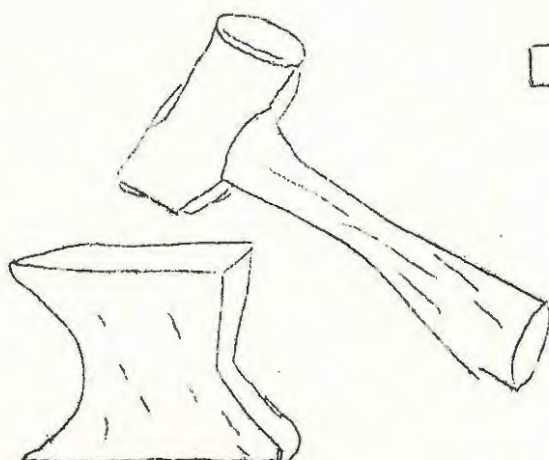
16962 "Anvil"
AFx3, AF2, AF2.



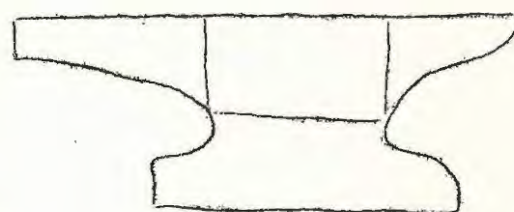
16663 "Anvil"
AFx3, AF2, AF1.



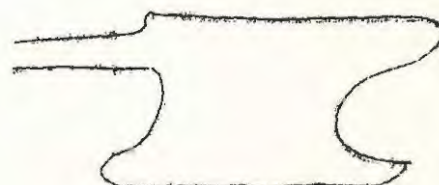
16248 "Anvil"
AFx3, AF2, AF2.



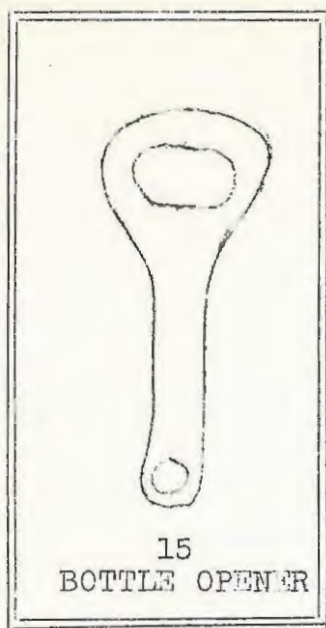
423 "Hammer & Anvil"
AFx2, AFo3, AFo3.



16221 "Anvil"
AFx3, AF2, AF2.

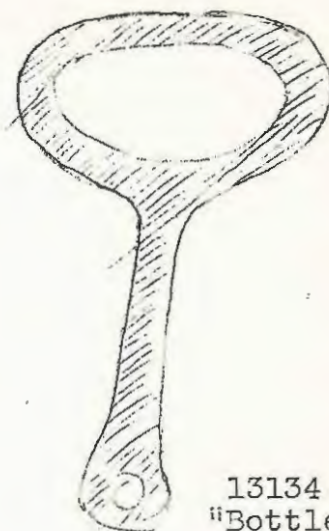
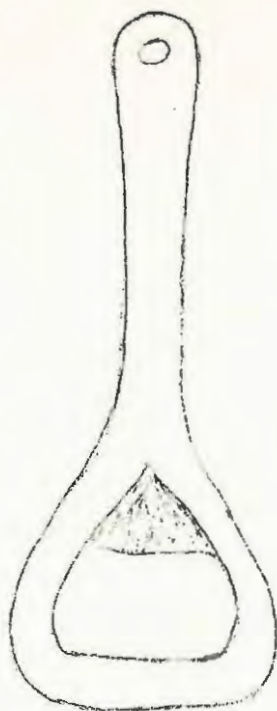


10441 "Anvil"
AFx3, AF2, AF2.

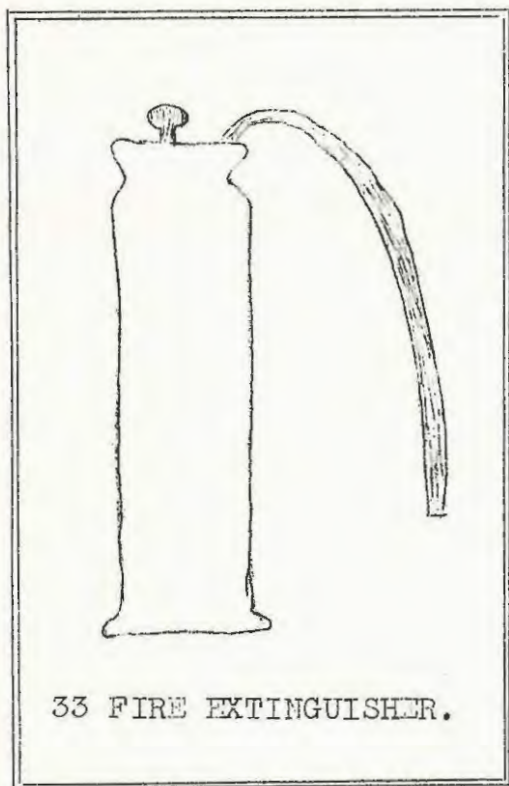


15
BOTTLE OPENER

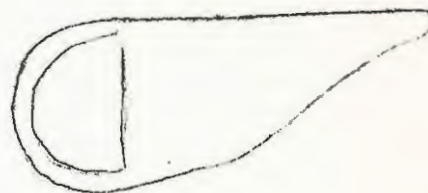
4531 "Bottle
opener"
AF3, AF3, AF3.



13134
"Bottle Opener"
AF3, AF3, AF3.



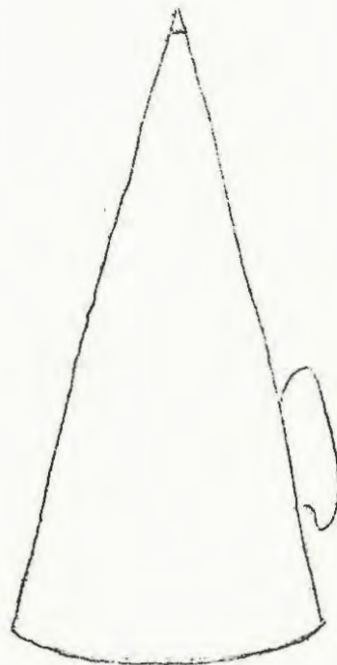
33 FIRE EXTINGUISHER.



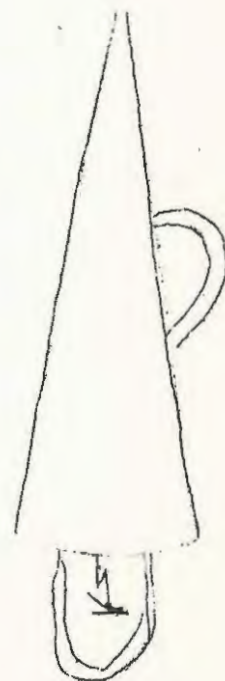
9764 "Bottle opener"
AFx2, ----- AFx3.



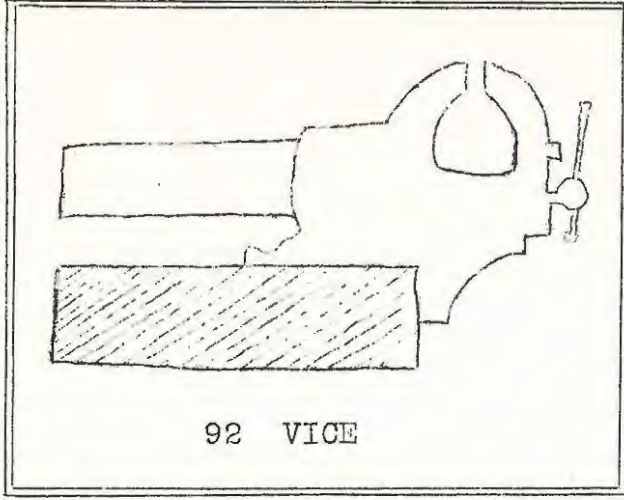
16120
"Fire
Extinguisher"
AFx1, AFo2,
AF2.



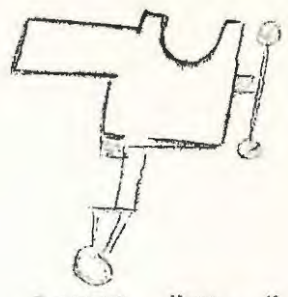
19699 "Fire
Extinguisher"
AFx1, AFo1, AF3.



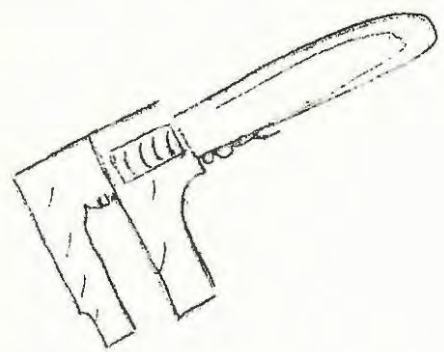
17674 "Fire
Extinguisher."
AFx1, AFo2, AF2.



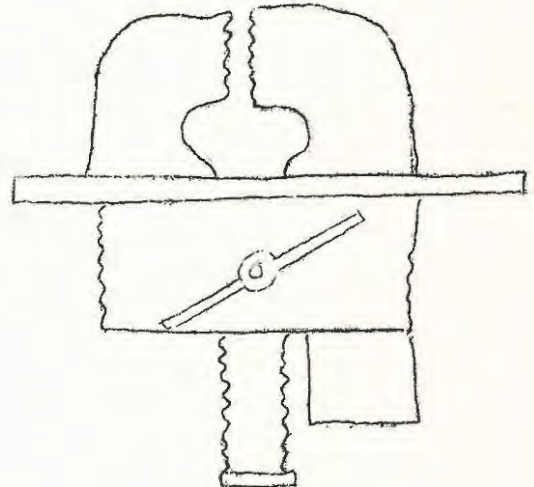
92 VICE



12738 "Vice"
AFo2, AF3, AF3.



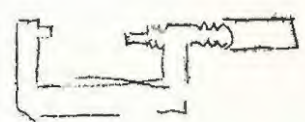
14179 "Spanner"
Hox2, H3, Hx3.



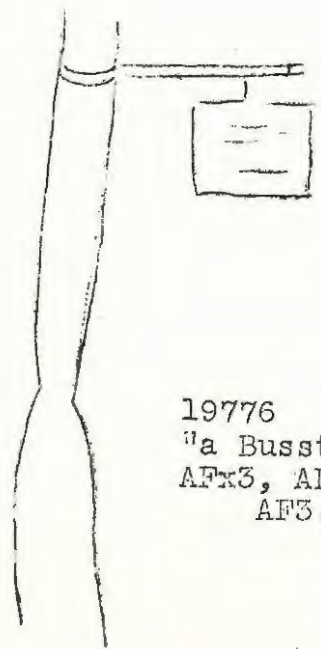
3003
"a VICE of sorts"
AF2o, AF3, AF3.



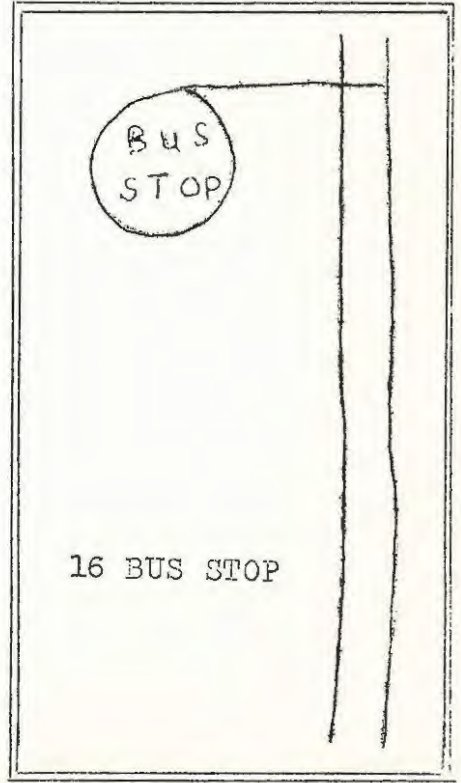
14702 "Saw" Hox1, Hx1, ----



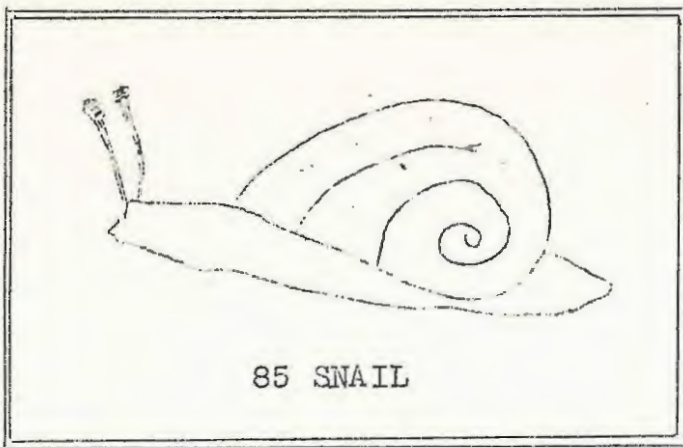
8479 "Micrometer Screw
Gauge" ---- Hx1, JH2.



19776
"a Busstop"
AFx3, AF2,
AF3.



16 BUS STOP

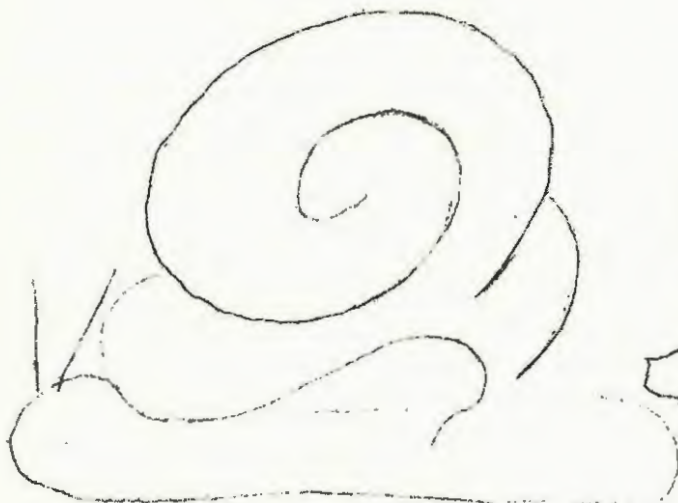


85 SNAIL



796 "Snail"
AF3, AF3, AF3.

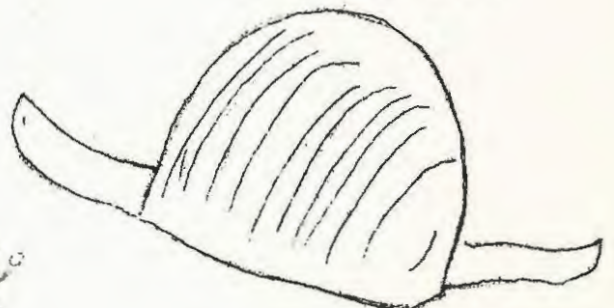
AF3, AF3, AF2.
19745 "Snail"



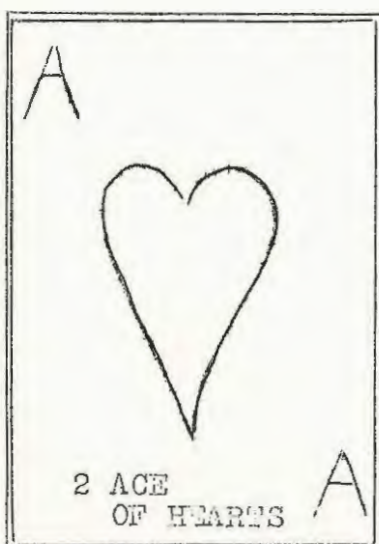
11811 "Snail" AF3, AF3, AF3.

8209 "Snail"
AF2, AF2, AF2,

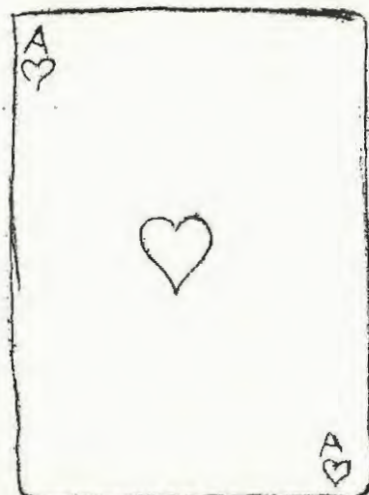
12656 "Snail" AF3, AF3, AF3.



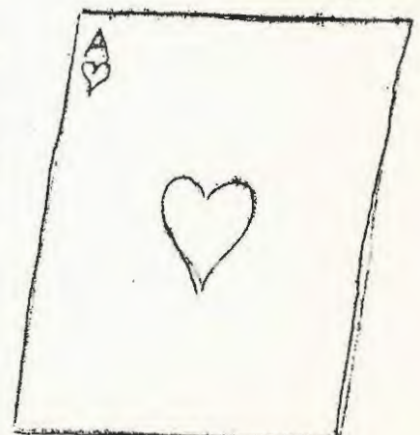
12935
"Ace of Hearts"
AFo3, AF3,
Afo3.



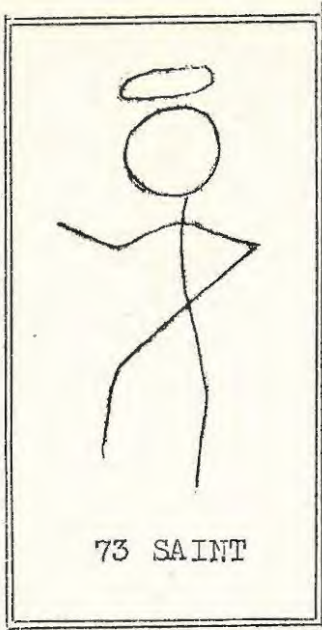
2 ACE
OF HEARTS



5645 "Ace of
Di Hearts"
AFo3, Af3, AFo3



2453 "Playing Card,
(Ace of Hearts)"
AHO3, AFo3, AFo3.



73 SAINT



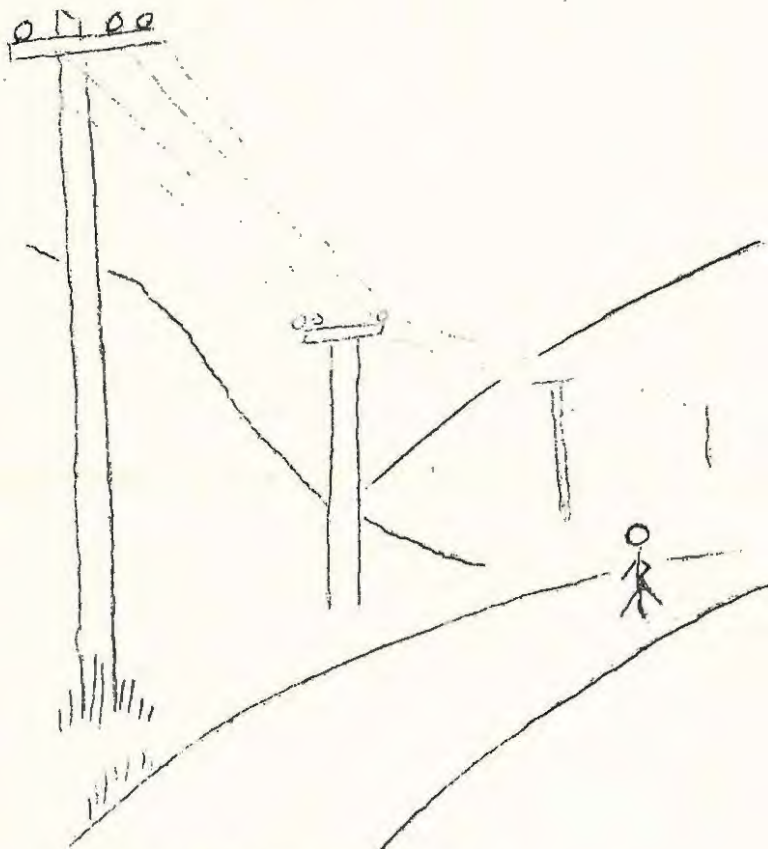
15677 "A Saint"
AF3, AF3, AF3.



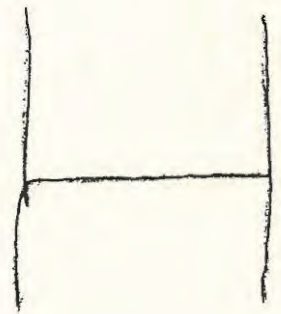
11498 "Saint"
AFox2, Ho2, AHox3



17361
"A Saint"
AF3
AF3
AF3



3503 "Telegraph poles"
Ho2 --- ---

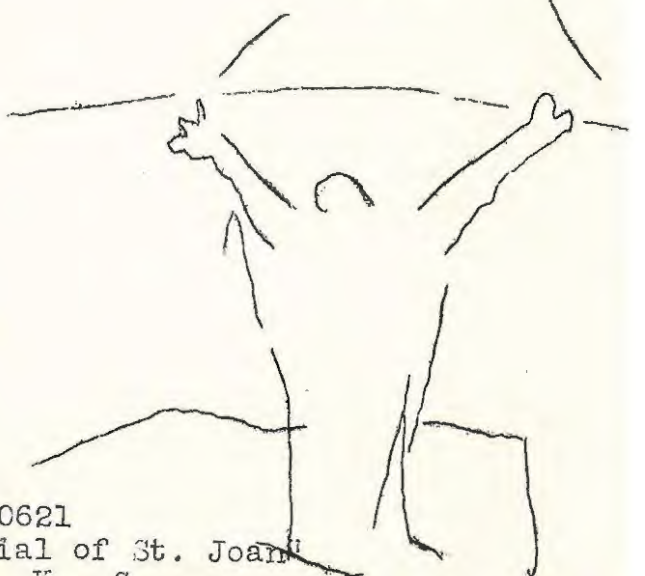


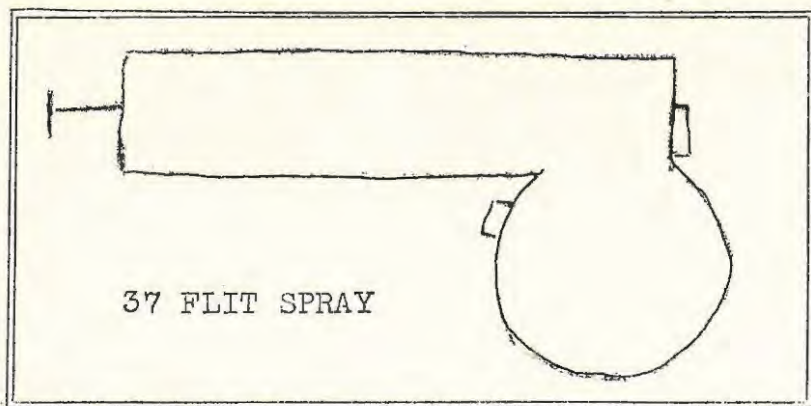
92 "Rugby game"
Hox1 --- ---

2112
Ho2
Ho2
Hox1
"Braaivleis"

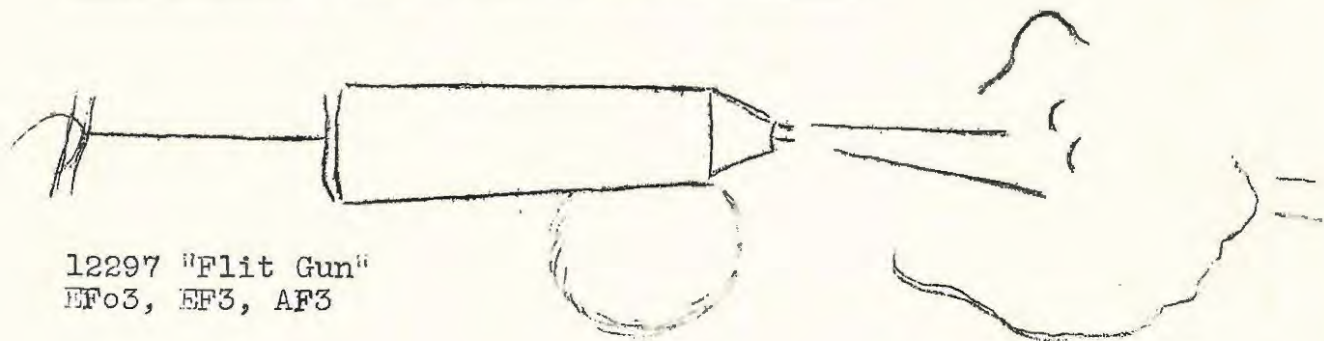


10621
"Trial of St. Joan"
A, K, C.

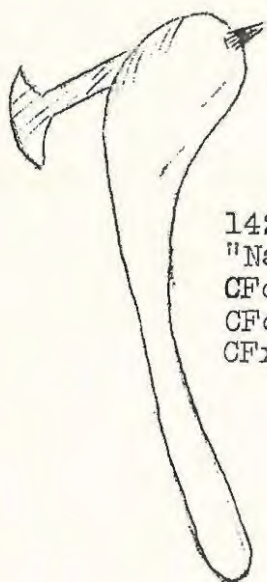




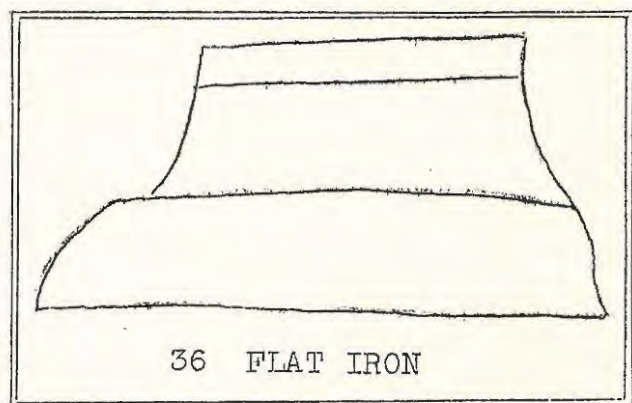
37 FLIT SPRAY



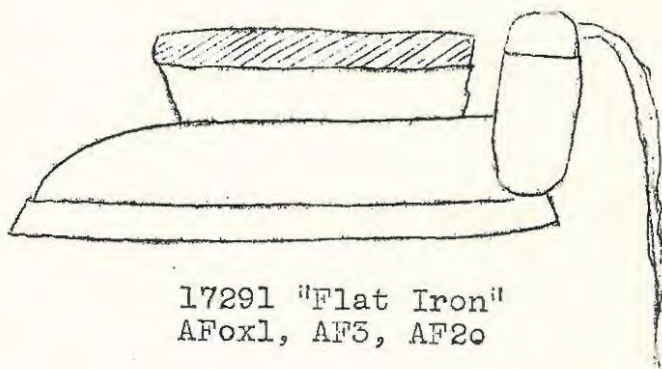
12297 "Flit Gun"
EFo3, EF3, AF3



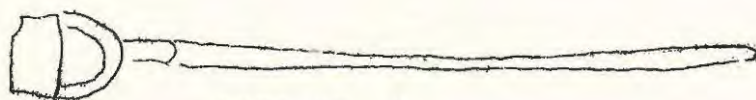
14233
"Native Hoe"
CFox1,
CFo3,
CFx2.



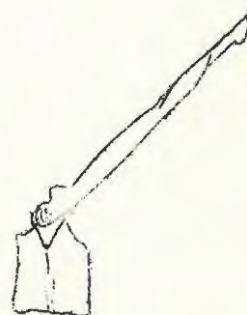
36 FLAT IRON



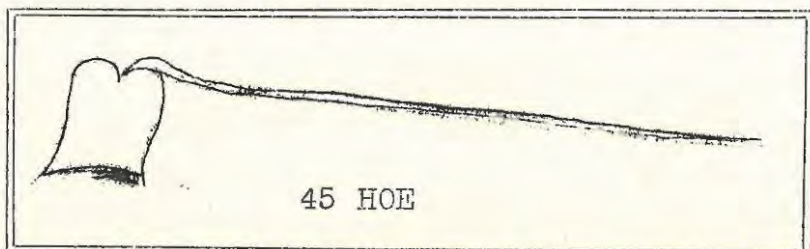
17291 "Flat Iron"
AFox1, AF3, AF2o



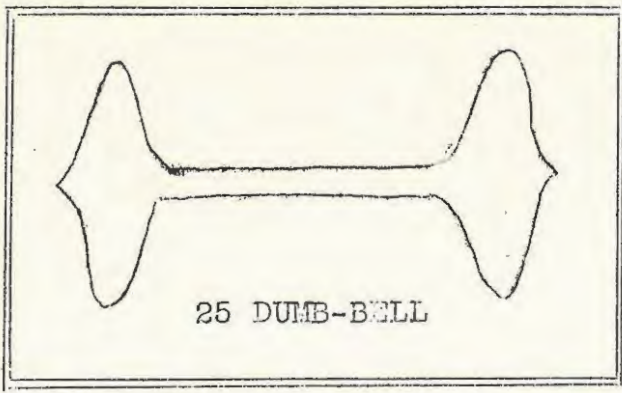
8572 "Hoe"
AFx2, AF2, AF3.



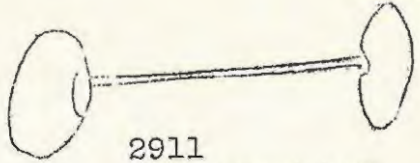
6515 "A hoe"
AF3, AF3, AF3.



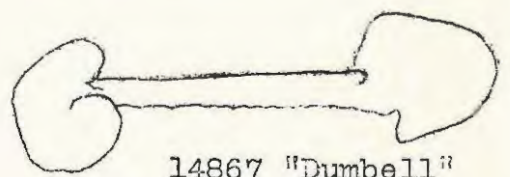
45 HOE



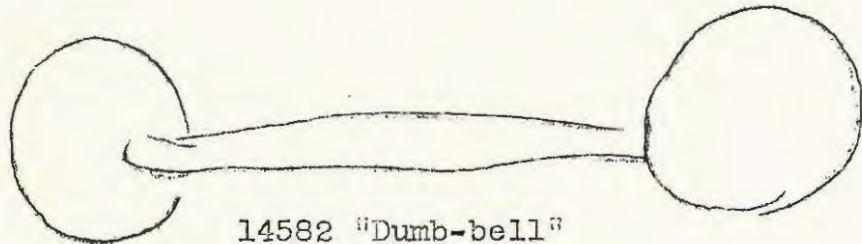
25 DUMB-BELL



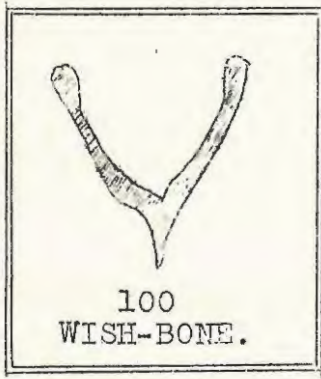
2911
"pr DUMBELLS"
AF3, AF2, AF3.



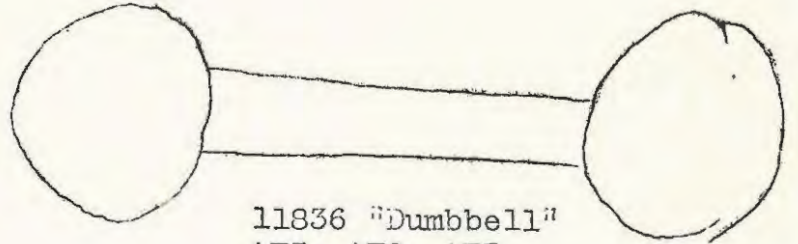
14867 "Dumbell"
AF2, AF2, AF3.



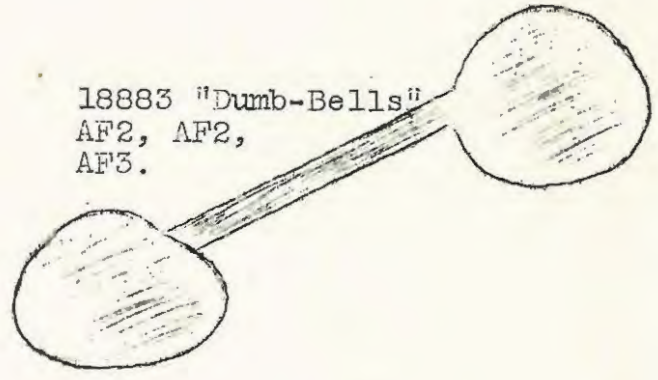
14582 "Dumb-bell"
AF2, AF2, AF3.



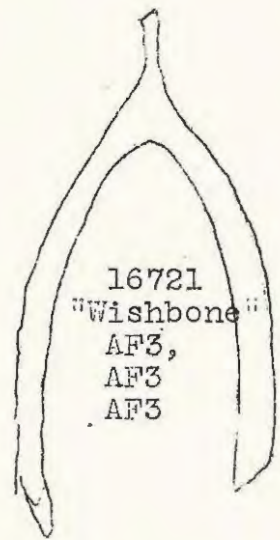
100
WISH-BONE.



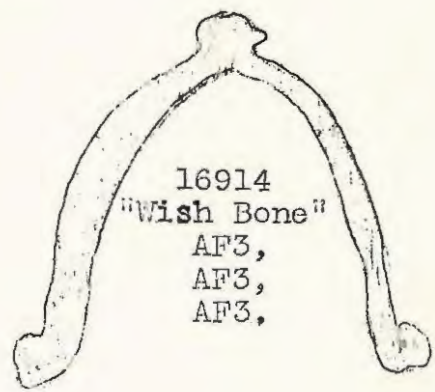
11836 "Dumbbell"
AF3, AF2, AF3,



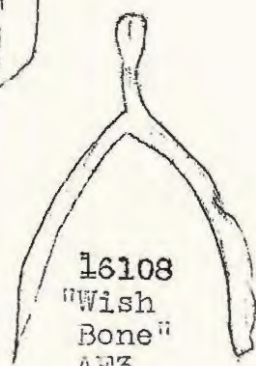
18883 "Dumb-Bells"
AF2, AF2,
AF3.



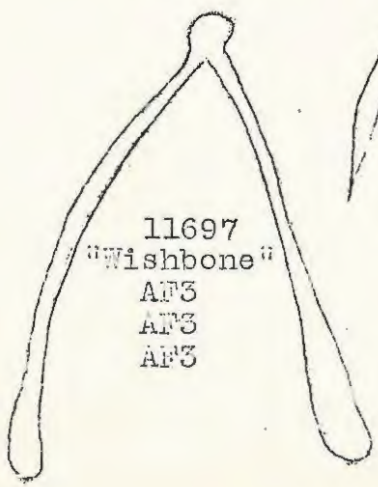
16721
"Wishbone"
AF3,
AF3
AF3



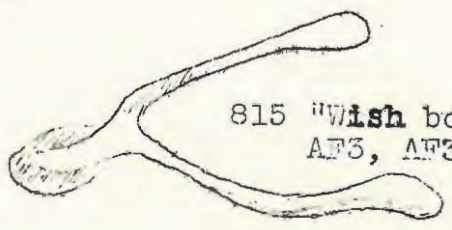
16914
"Wish Bone"
AF3,
AF3,
AF3.



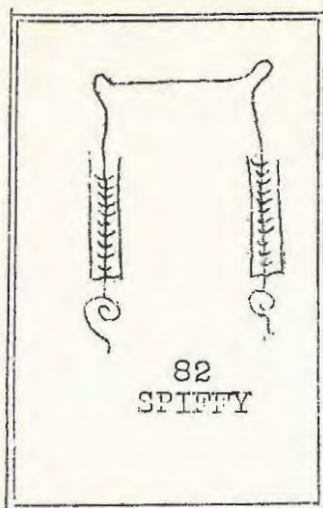
16108
"Wish
Bone"
AF3
AF3
AF3



11697
"Wishbone"
AF3
AF3
AF3



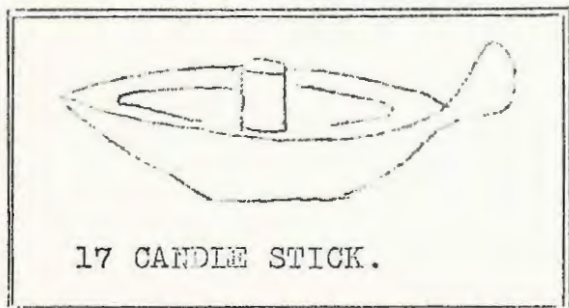
815 "Wish bone"
AF3, AF3, AF3.



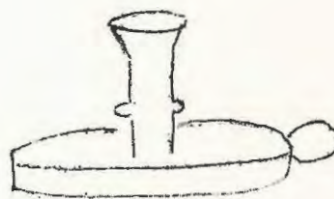
82
SPIFFY



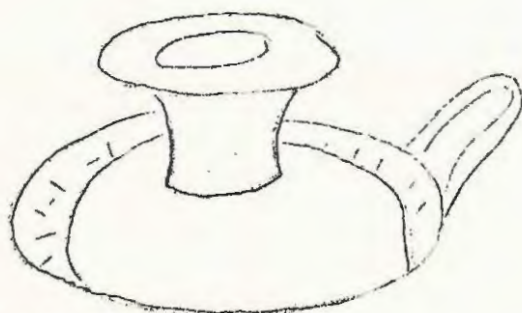
10500 "Spiffy"
AF3, AF2, AF3.



17 CANDLE STICK.



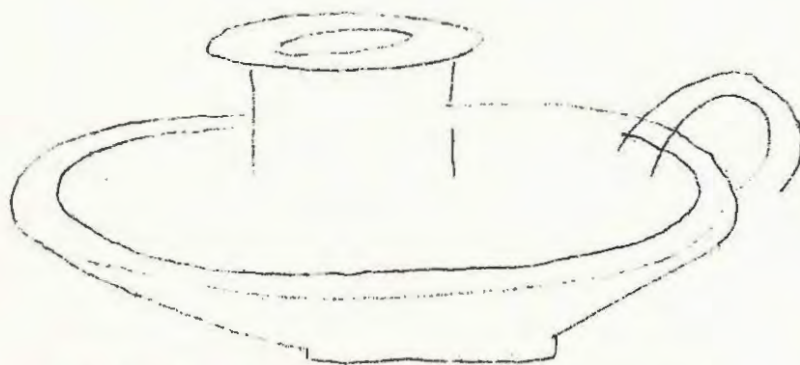
2584 "Candlestick"
AFx2, AF2, AF3.



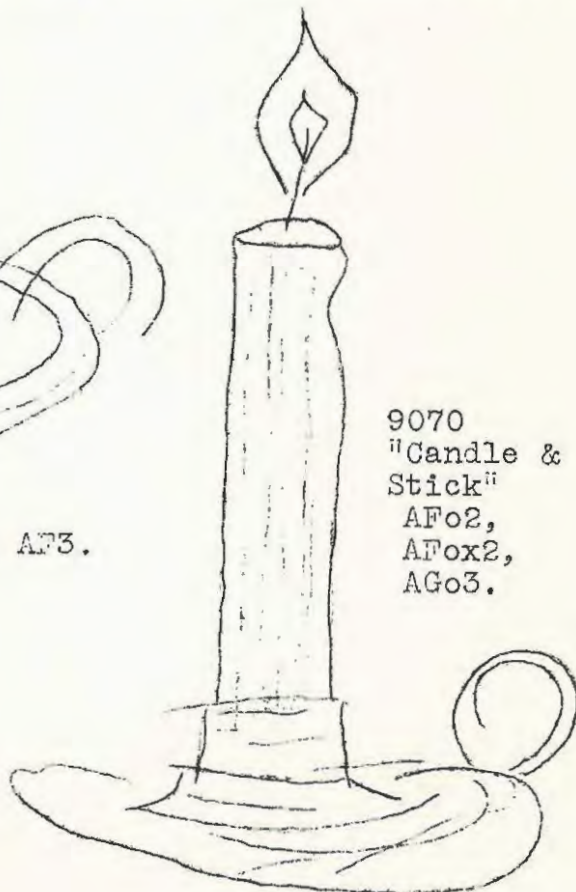
1603 "Candlestick"
AF2, AF2, AF3.



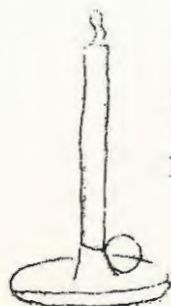
2783 "Candlestick" AFx1, AF2, AF2.



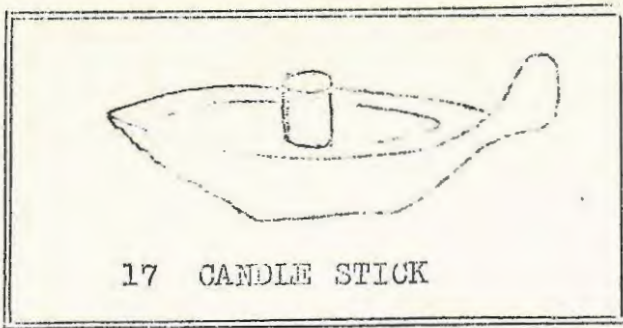
9364 "Candle holder" AF3, AF2, AF3.



9070
"Candle &
Stick"
AFo2,
AFox2,
AGo3.

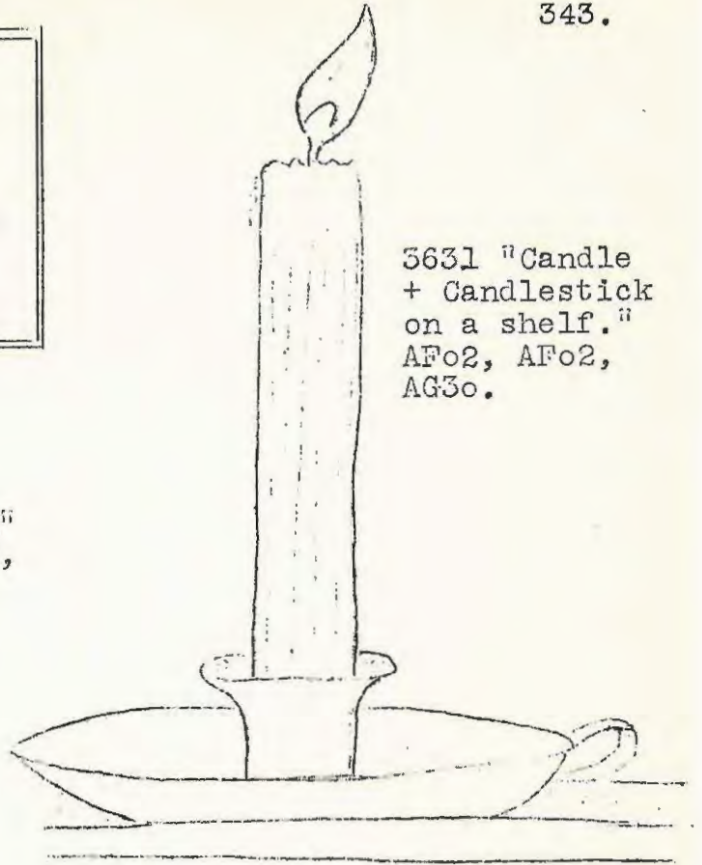


2323 "Candle and
Candle-stick"
Fox1, AFo2, Go2

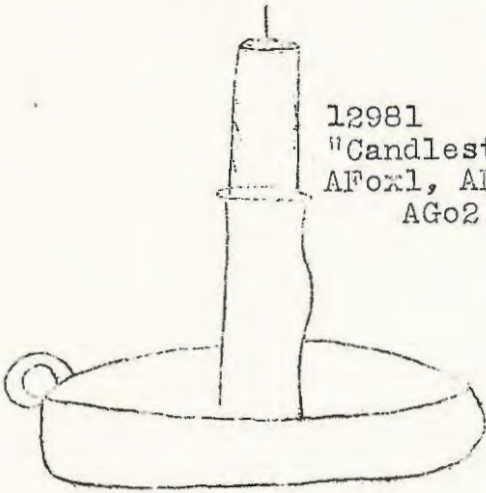


17 CANDLE STICK

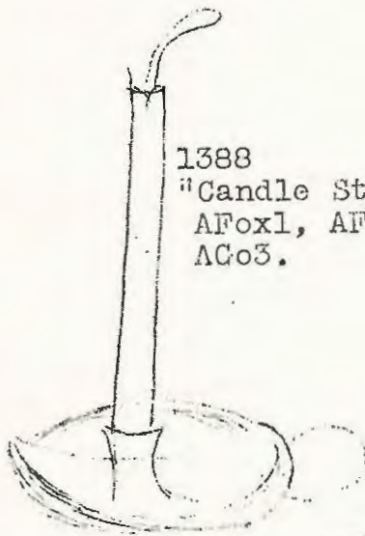
3631 "Candle + Candlestick on a shelf."
AFo2, AFo2, AG3o.



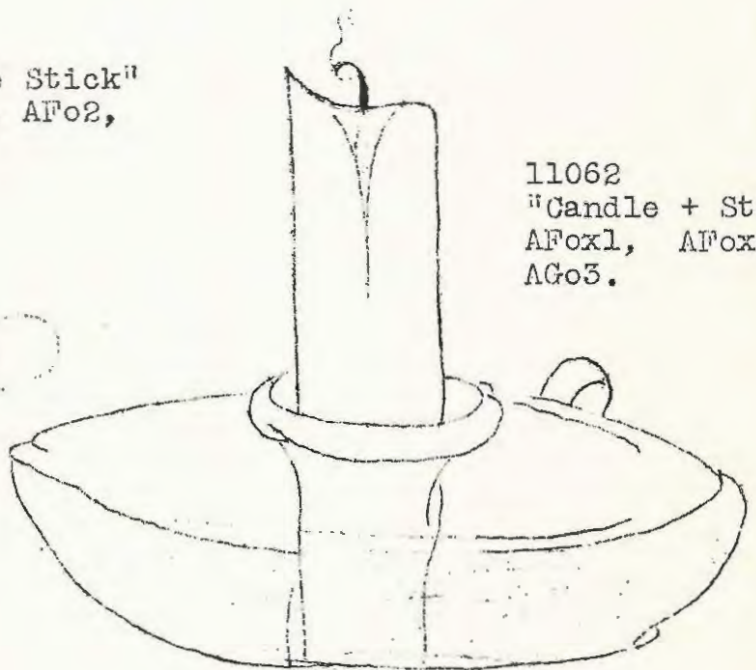
12981
"Candlestick"
AFox1, AFox2, AGo2.



1388
"Candle Stick"
AFox1, AFo2, AGo3.



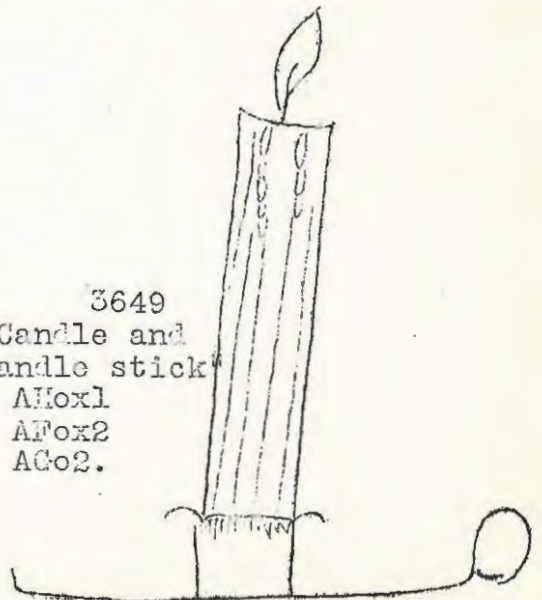
11062
"Candle + Stick"
AFox1, AFox2, AGo3.

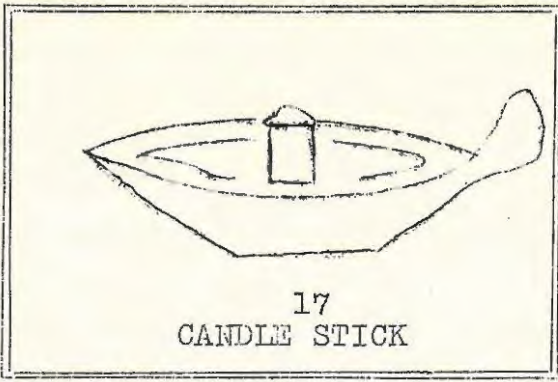


2497 "Candle in Candlestick"
AFox2, AFo2, AGo2



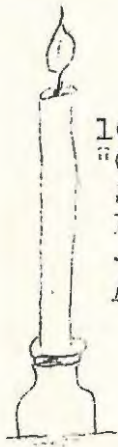
3649
"Candle and Candle stick"
AHox1
AFox2
AGo2.





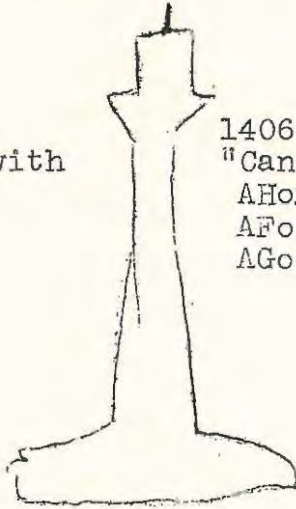
17
CANDLE STICK

5644
"Candle in
Stick"
AHox2
AFo2
AGo2

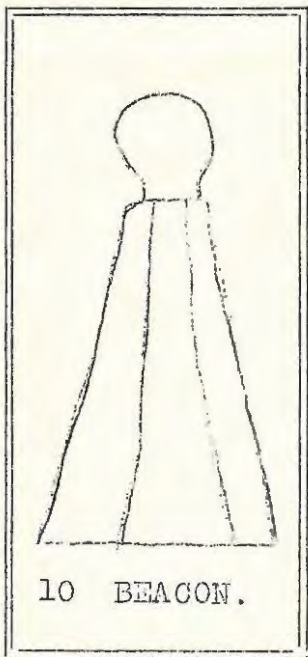
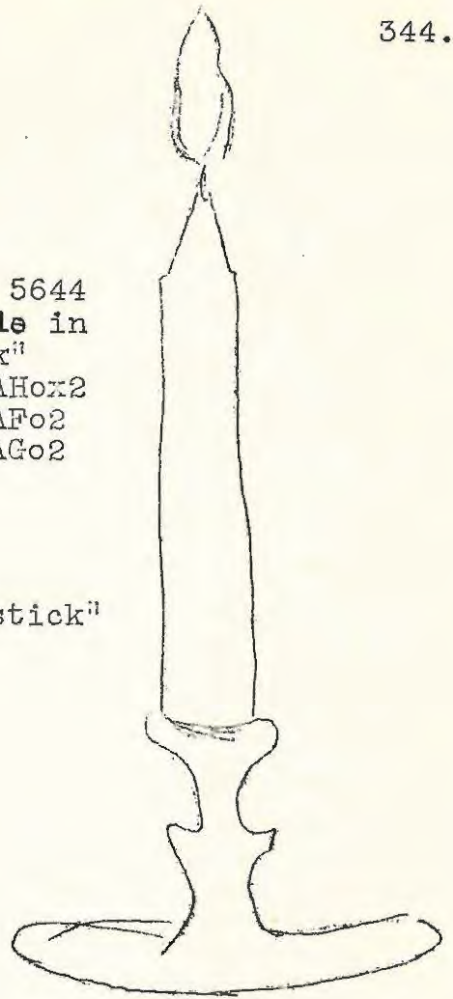


10765
"Candle with
Stand"
EHox1

AGo1



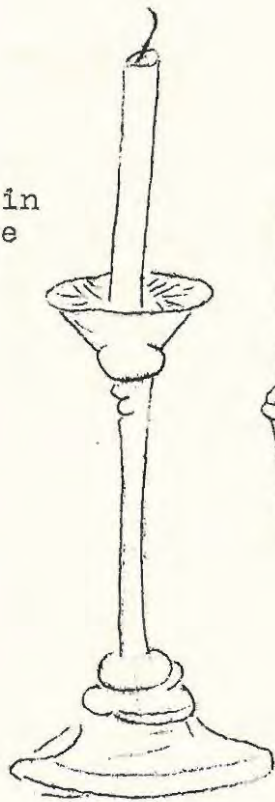
14066
"Candlestick"
AHox1
AFo2
AGo1



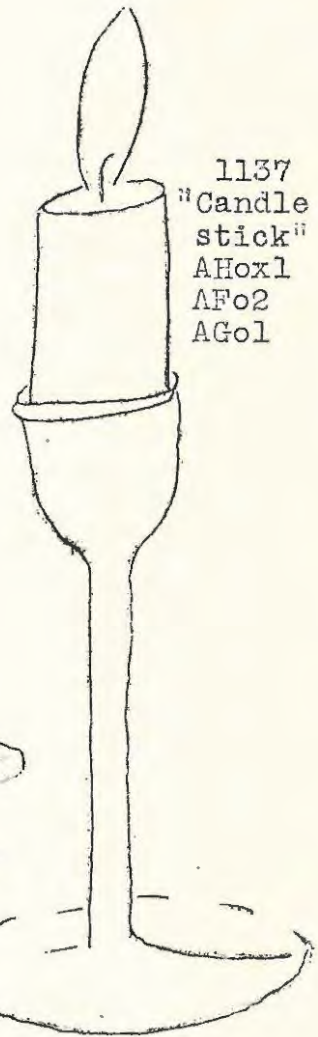
10 BEACON.



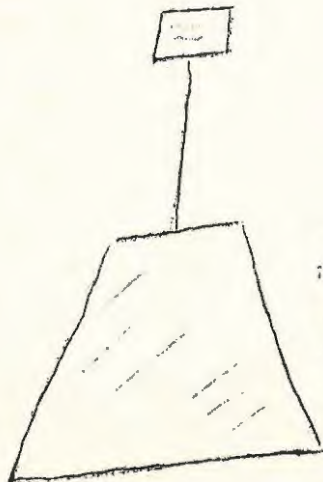
19285
"Candle in
elaborate
c-stick"
AHox1
AFo2
AGo1



270
"Candle stick
and Candle"
AHox1
AFo2
AGo1.



1137
"Candle
stick"
AHox1
AFo2
AGo1



3194
"Beacon"
AF2
AF2
AF1.



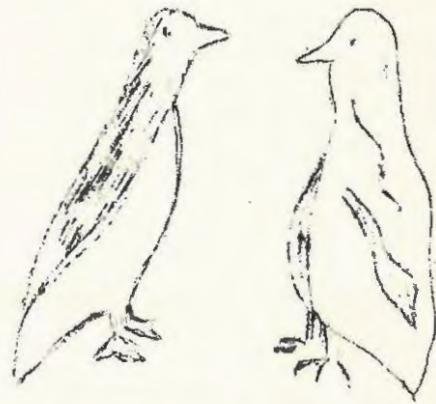
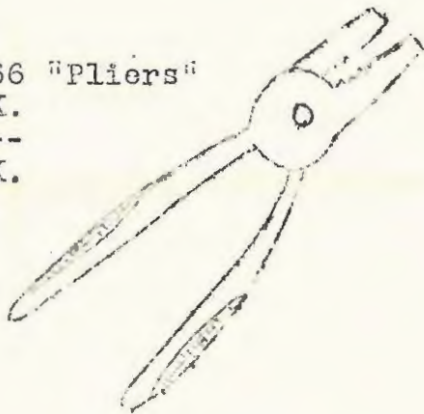
60 PENGUIN.



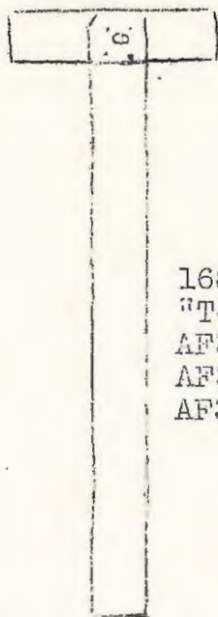
13681 "Penguin"
AF3, AF3, AF3.

10966 "Pliers"
60 K.

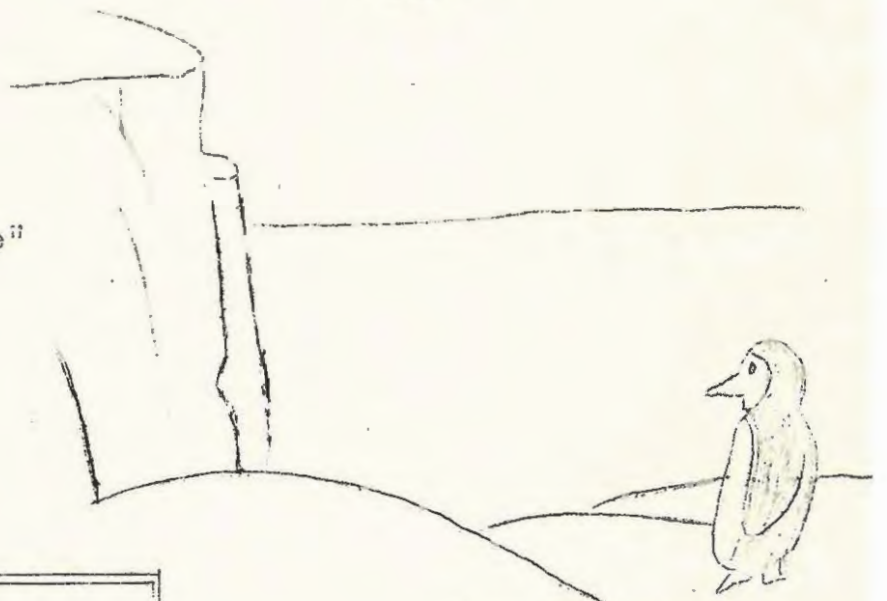
60 K.



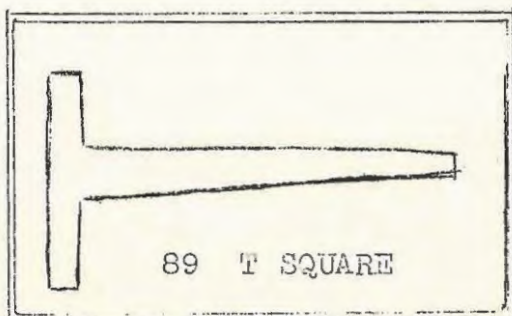
8388 "Penguin talk"
AF3, AF3, AF3.



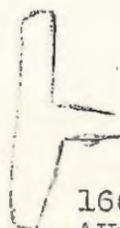
1682
"T-Square"
AF3
AF3
AF3



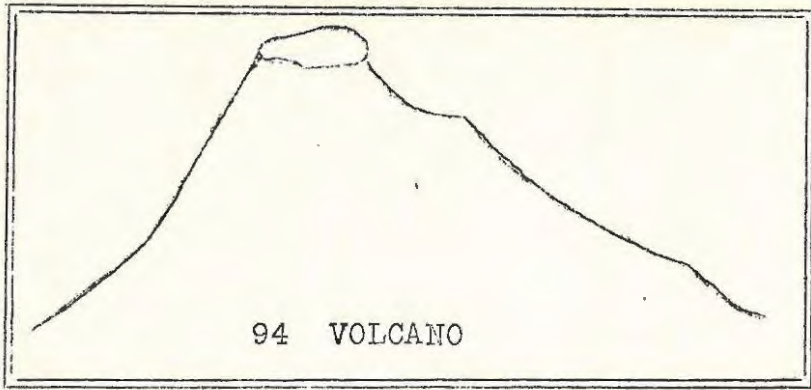
5720 "Moonlit sea
and penguin"
AF2
AF3
AF3



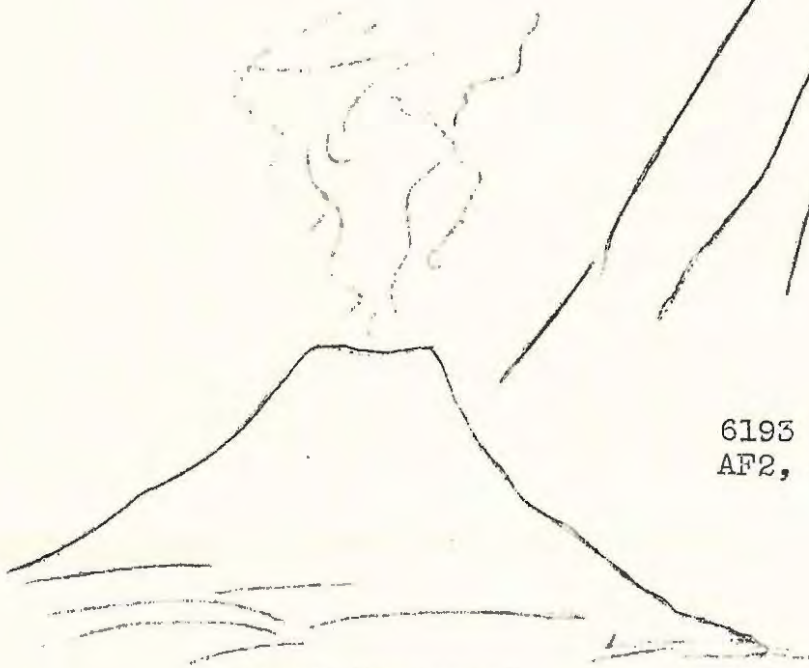
89 T SQUARE



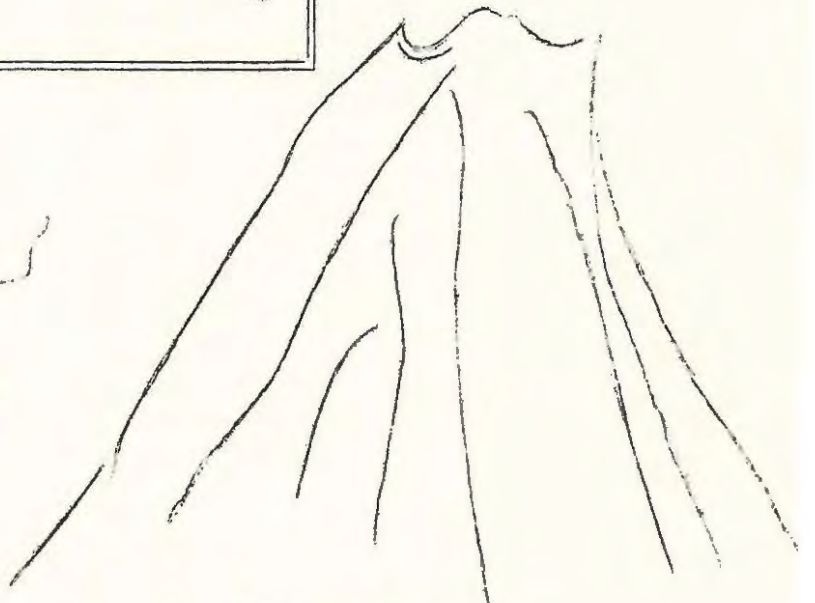
16664 "Broken 'T' Square"
AHx3, AH3, AE3.



94 VOLCANO



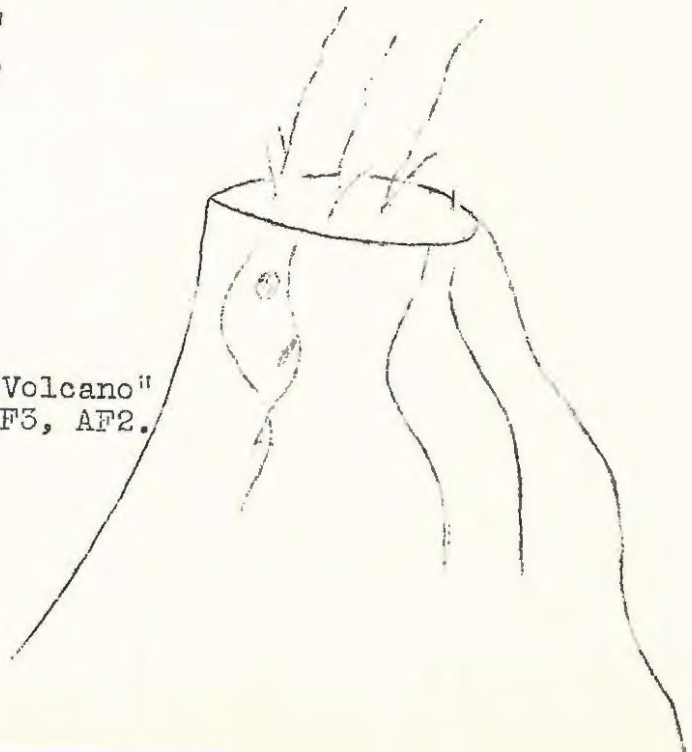
14698 "Volcano"
AFo3, AF3, AF3,



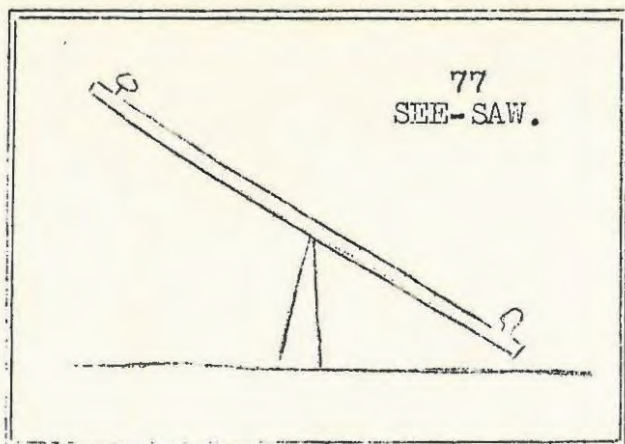
6193 "Volcano"
AF2, AF3, AF2.



4713 "Volcano"
AF3, AF3, AF3.



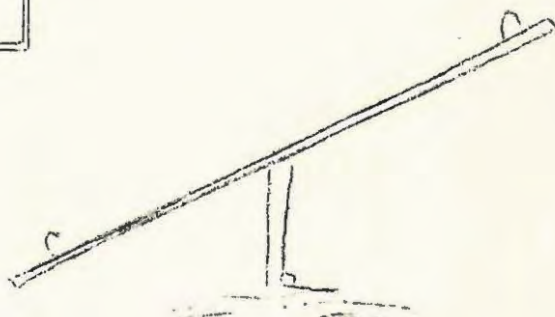
12265 "Volcano"
AFo3, AF3, AF2.



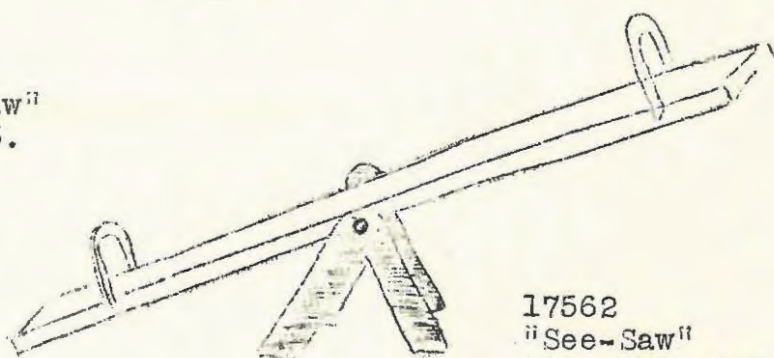
3570 "Open Book"
F01 --- ---



11829 "See-Saw"
AF3, AF3, AF3.



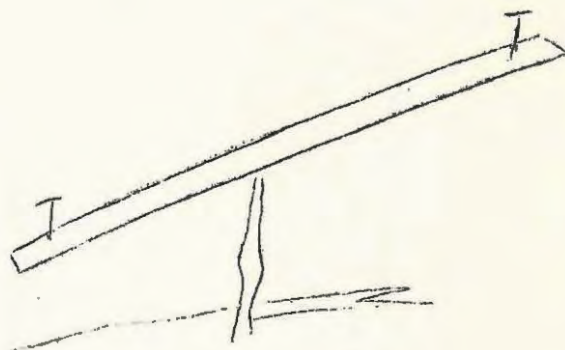
6359 "See Saw"
AF3, AF3, AF3.



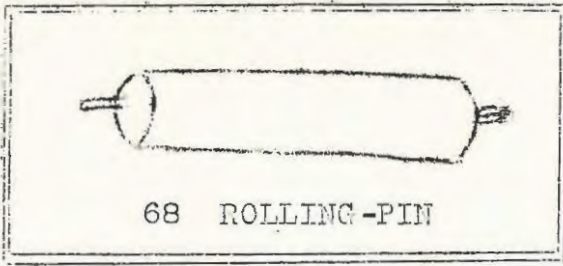
17562
"See-Saw"
AF3, AF3, AF3.



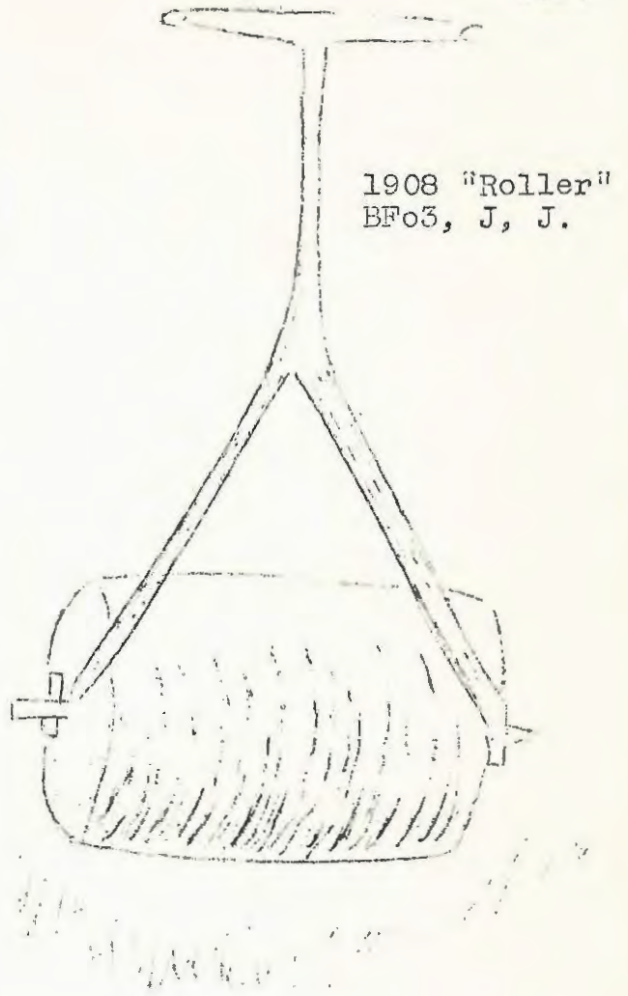
11040 "Balance"
Ho3J, ---- Fox3.



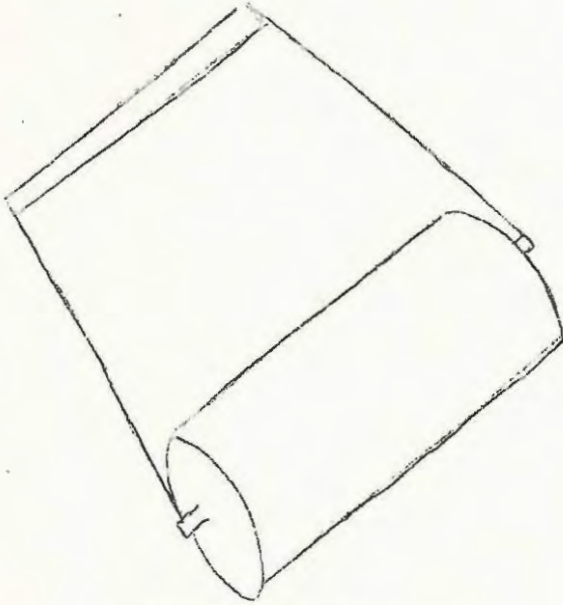
12942 "See-Saw"
AF3, AF3, AF3.



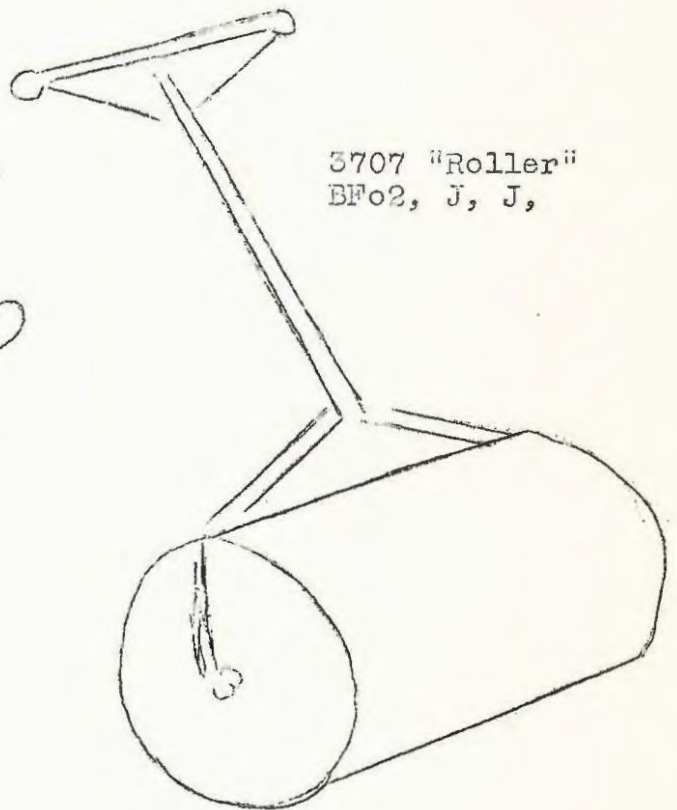
68 ROLLING-PIN



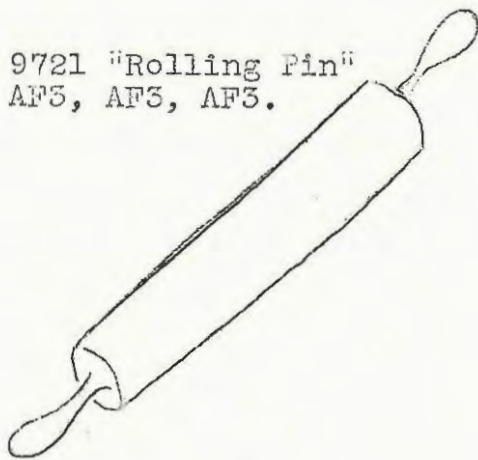
1908 "Roller"
BFo3, J, J.



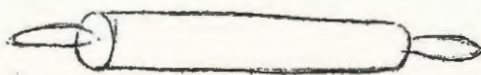
2353 "Roller"
BFo3, J, J.



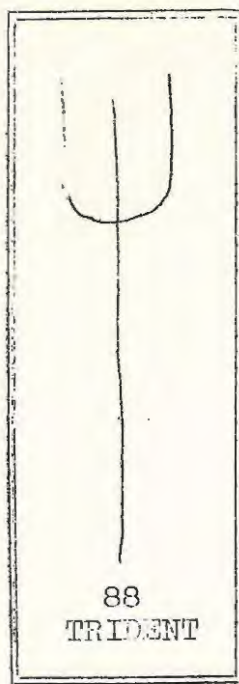
3707 "Roller"
BFo2, J, J,



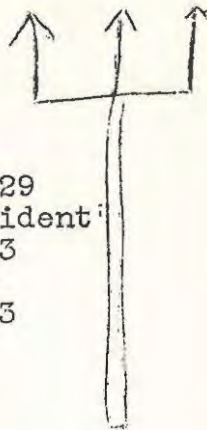
9721 "Rolling Pin"
AF3, AF3, AF3.



3358 "Rolling Pin"
AF3, AF3, AF3.



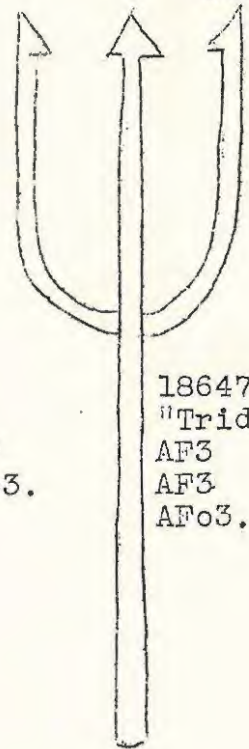
88
TRIDENT



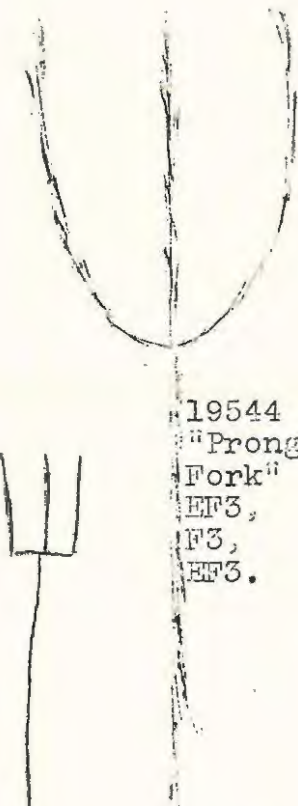
14829
"Trident"
AFo3
AF3
AFo3



10031 "Fork"
Fo2, F3, BFo3.



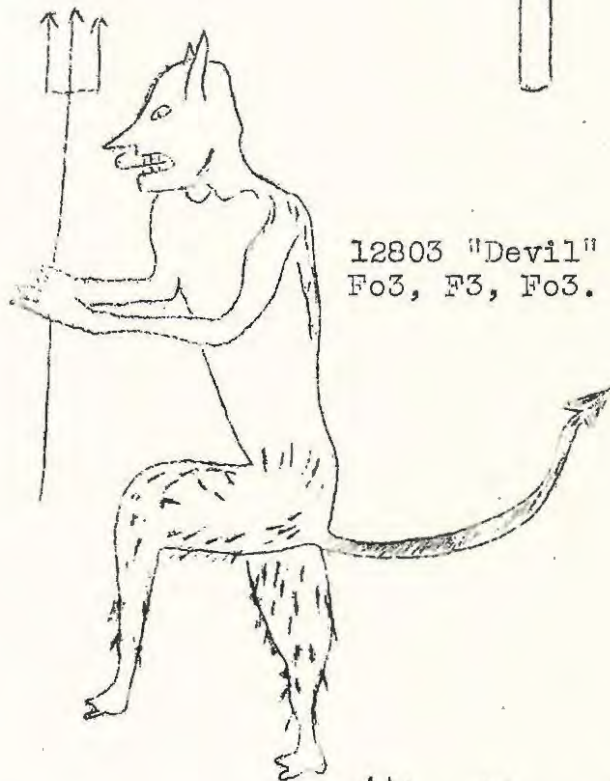
18647
"Trident"
AF3
AF3
AFo3.



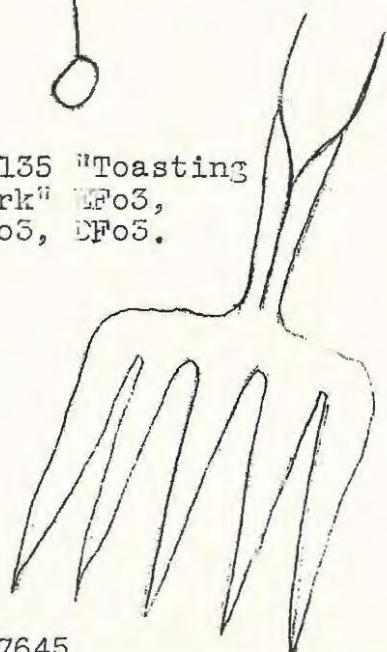
19544
"Pronged
Fork"
EF3,
F3,
EF3.



12135 "Toasting
Fork" EFo3,
EFo3, DFo3.



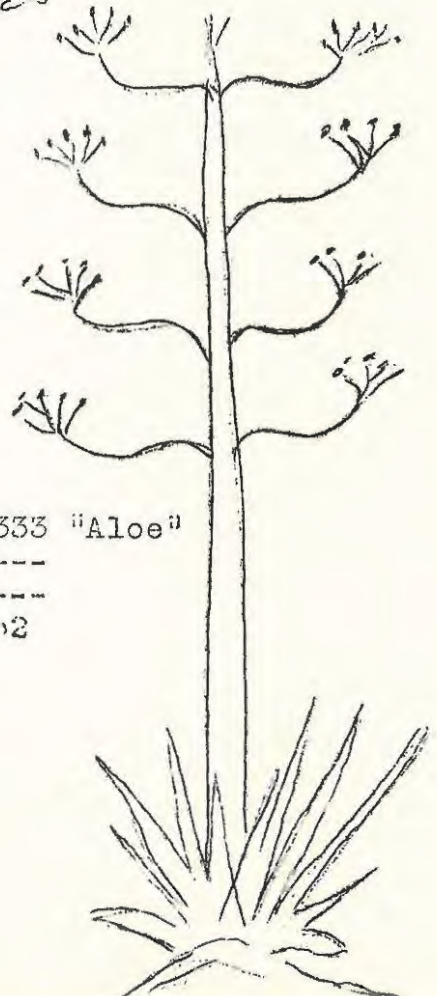
12803 "Devil"
Fo3, F3, Fo3.



7645
"Gardening Fork"
---- BFox2, BFox2.

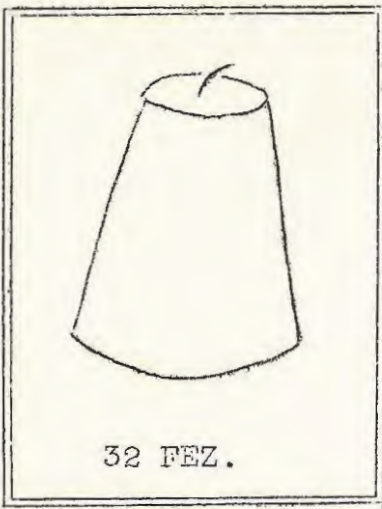


12333
"Toasting
Fork"
EFo3
EFo3
DFo3

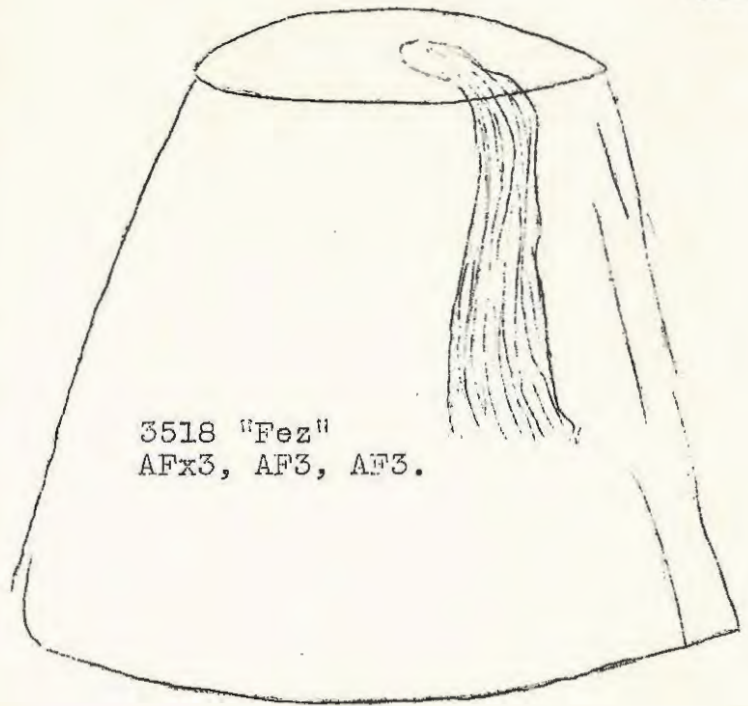


6333 "Aloe"

Co2



32 FEZ.



3518 "Fez"
AFx3, AF3, AF3.



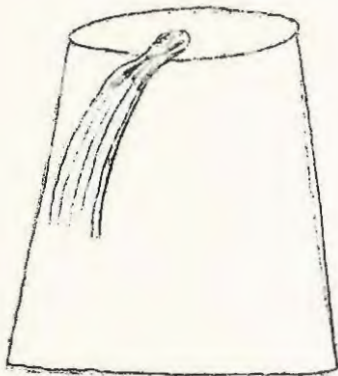
4066
"Fez on a Malay"
AFo3, AF3, AF3.



19696
"Fez"
AF3, AF3
AF3.



35
"Malay"
Fo3
Fo3
Fo3



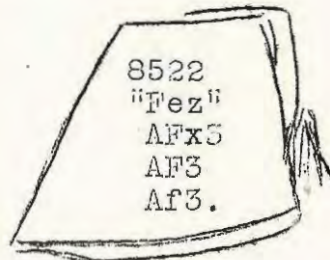
7382 "A Mohammedan's hat (I think
a badsi) fez" AF3x, AF3, AF3.



11093 "Fez"
AFx3, AF3, AF3.



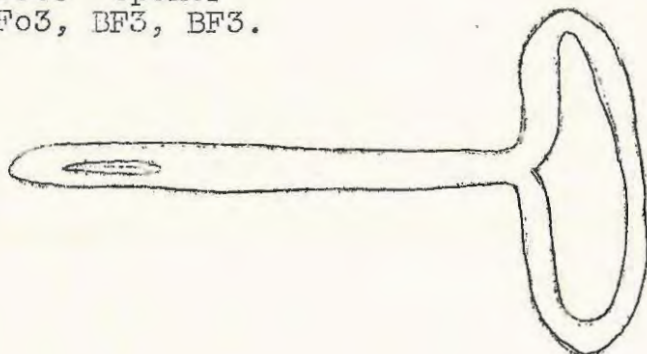
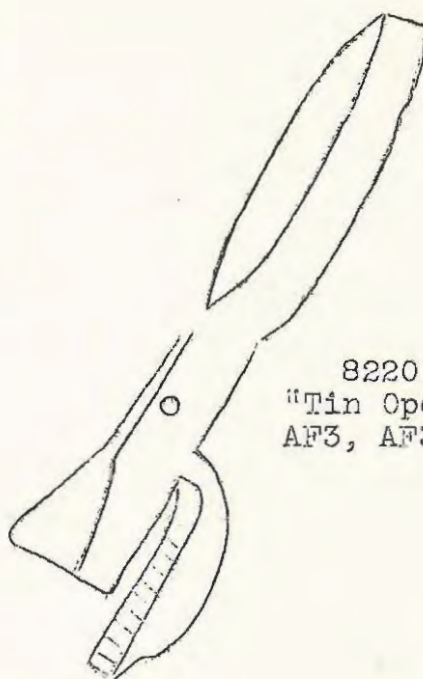
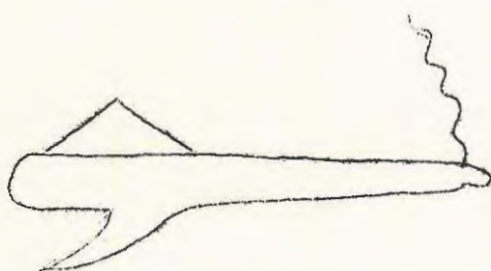
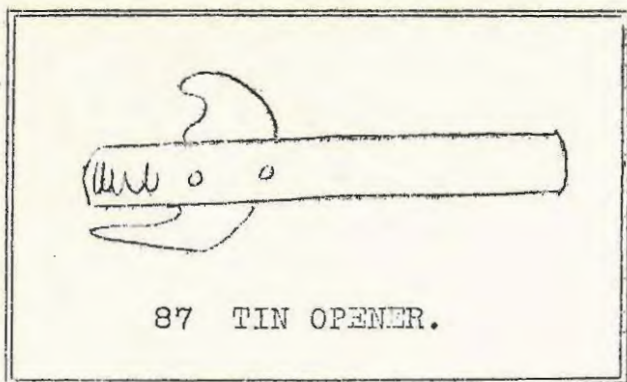
13723 "Organ Grinder's Monkey"
Fo3. Fx2. Fo3.

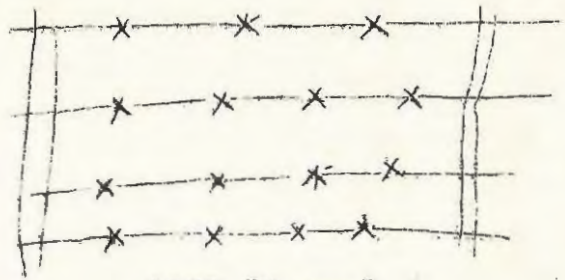
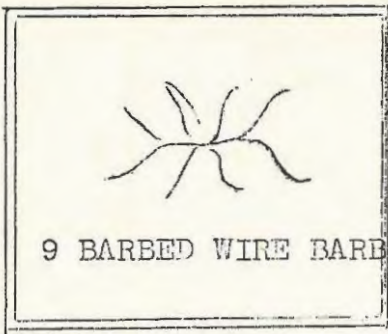


8522
"Fez"
AFx3
AF3
AF3.



17198
"Gyppo"
Fo3
Fx2, Fo3.



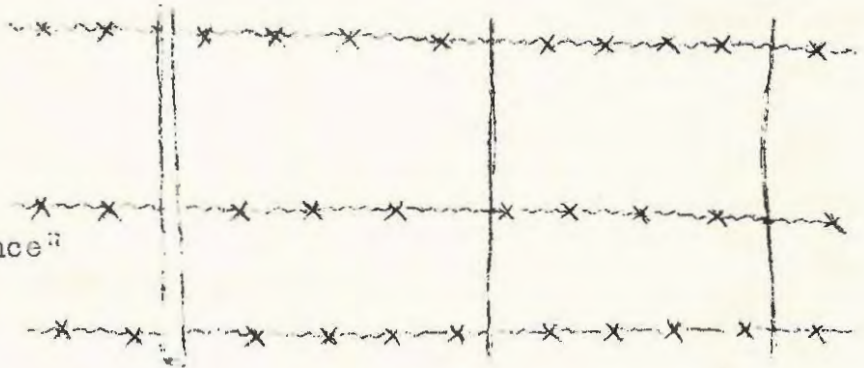


8253 "Fence"
Go2, Fo3, Fo3.



9374 "A Fence"
Go2, Fo3, Fo3.

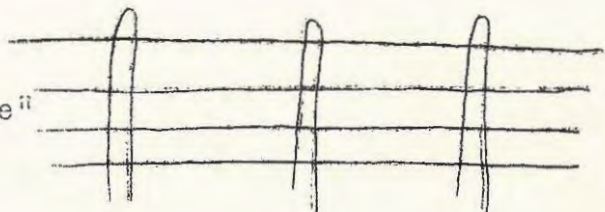
12368
"Barbed Wire Fence"
Go2, Fo3, Fo3.

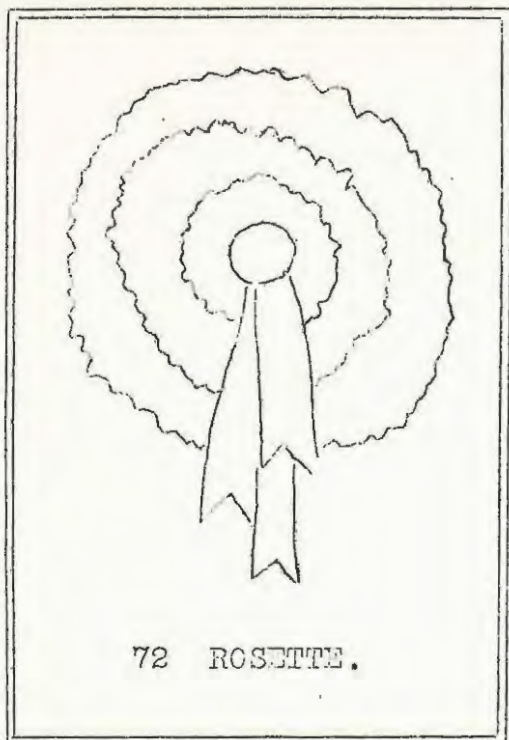


4476 "Fence"
G2, Fo3, Fo3.

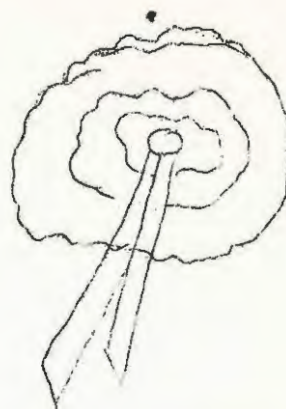


4351 "Fence"
I, ---, I.

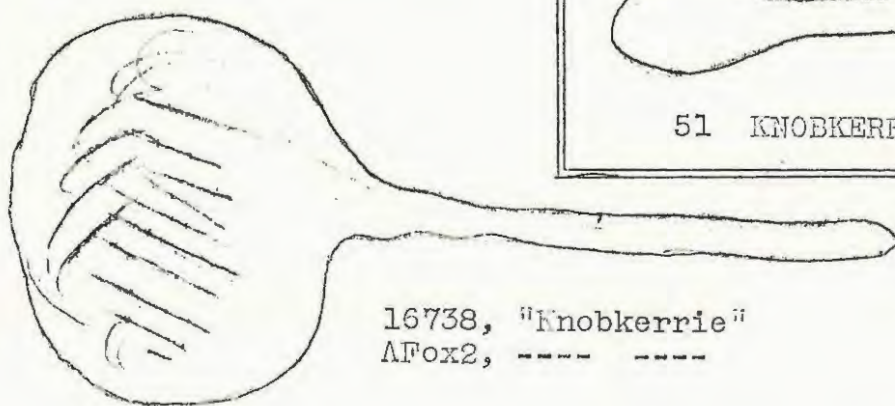




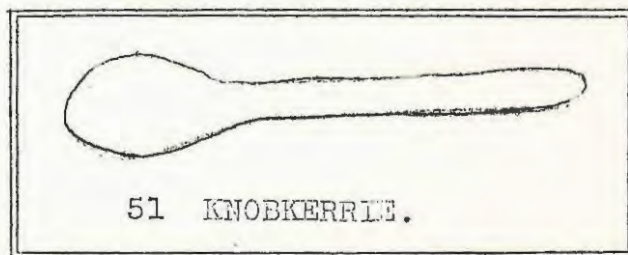
72 ROSETTE.



11190 "Rosette"
AF3, AF3, AF5.



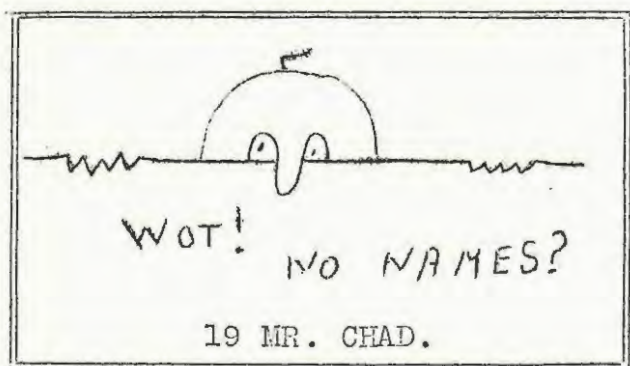
16738, "Knobkerrie"
AFox2, ---- ----



51 KNOBKERRIE.

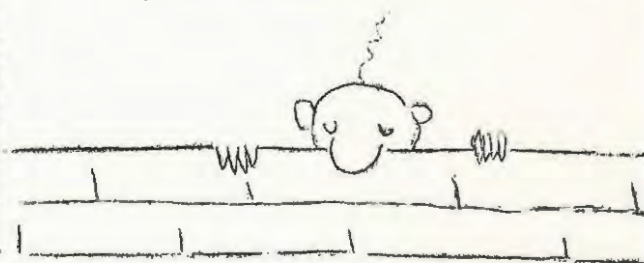


18091
"Native Knobkerrie"
AF3, AF3, CF3.



WOT!
NO NAMES?

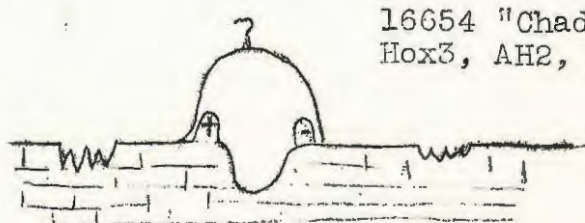
19 MR. CHAD.



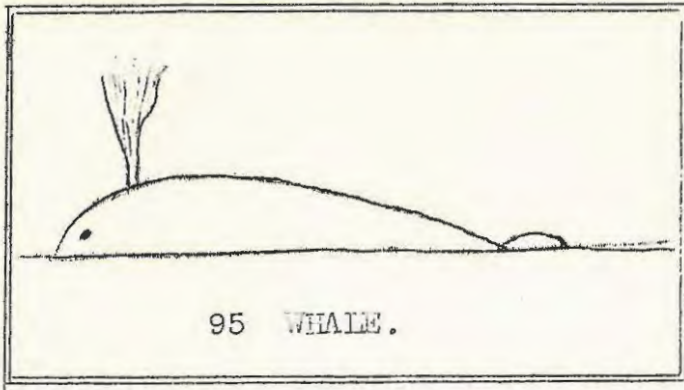
7255. "Brick wall Chad leaning
over"
Hox3, AHo2, AH3.



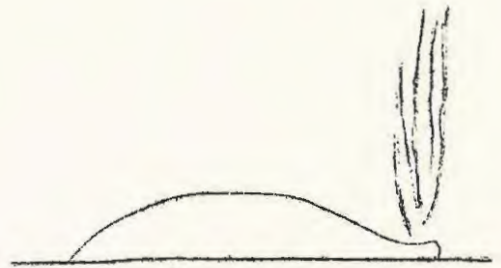
19617 Unnamed.
Hx3, AHo2, AH3.



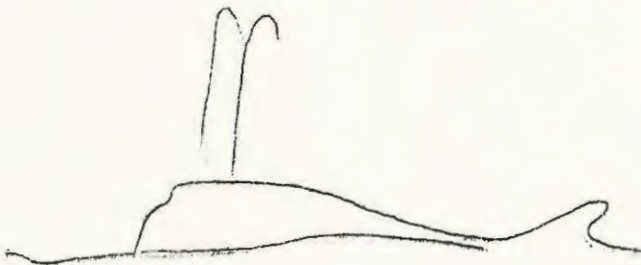
16654 "Chad"
Hox3, AH2, AH3.



95 WHALE.



4048 "Whale"
AF3, AF3, AF3.



8692 "Whale" AF3, AF3, AF3.



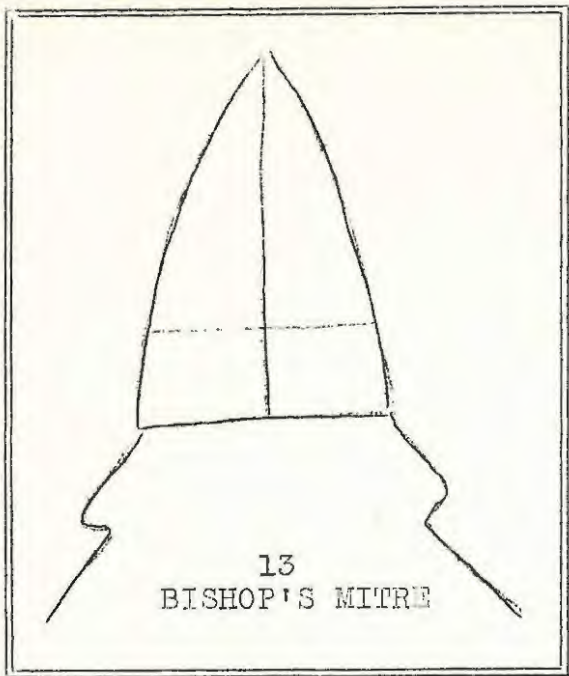
9480 "Whale" AF2, AG2, AH3.



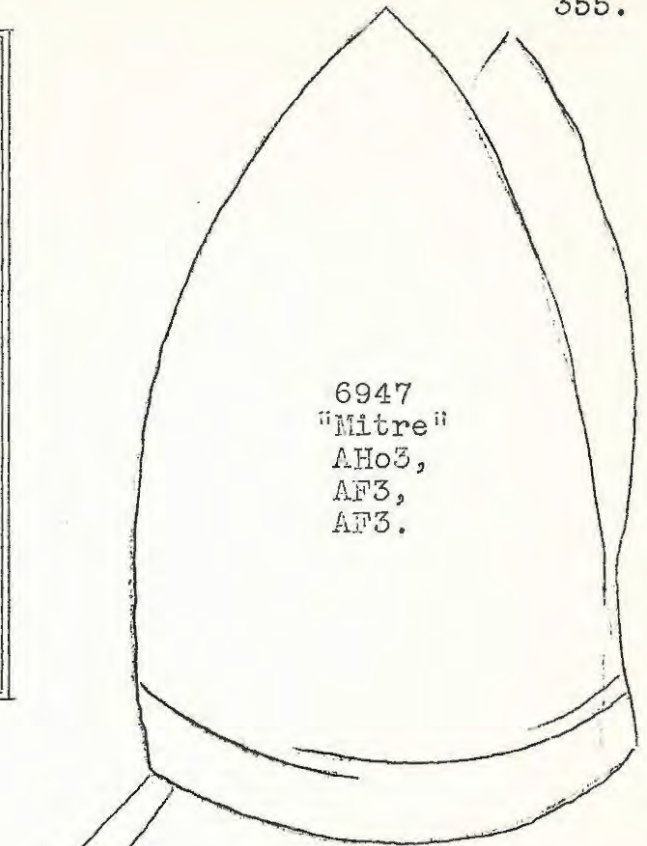
12674 "Whale" AF1, AG2, AH3.



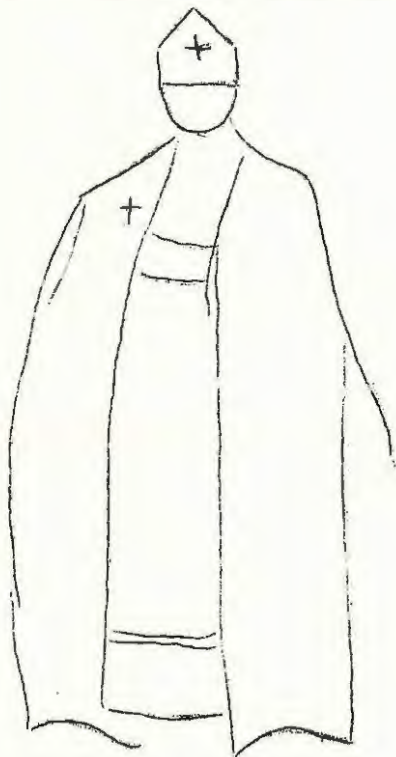
16912. "Whale" AF2, AG2, AF3.



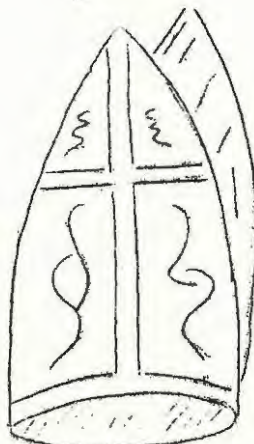
13
BISHOP'S MITRE



6947
"Mitre"
AHo3,
AF3,
AF3.

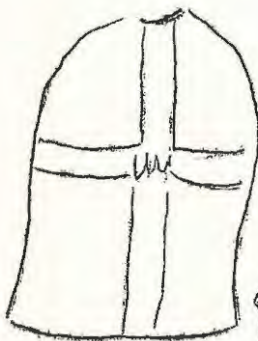


7311 "Bishop"
F3, Fox2, Fo3.

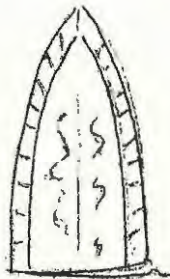


8326 "Mitre"
AFo3, AF3, AF3.

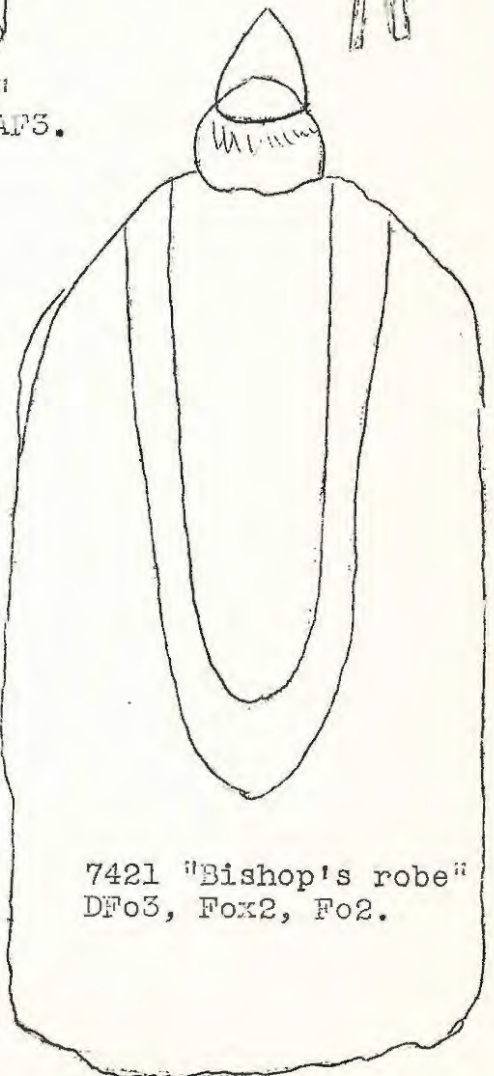
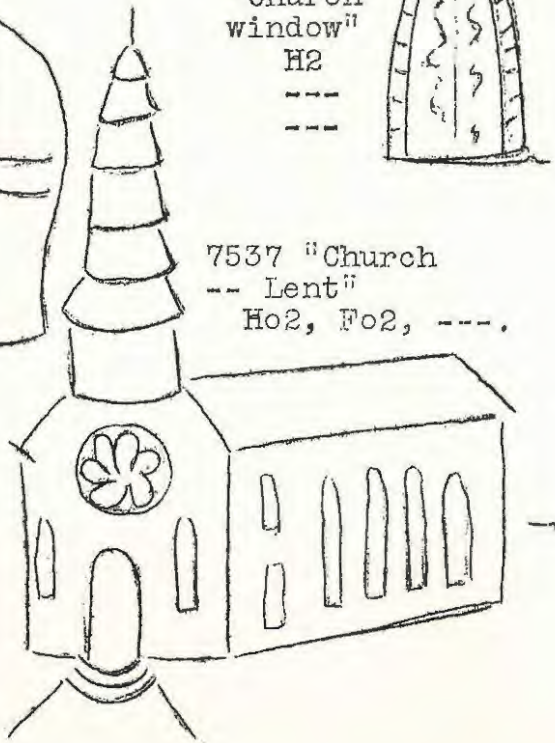
6111
"Bishop's
Mitre"
AF3, AF3
AF3.



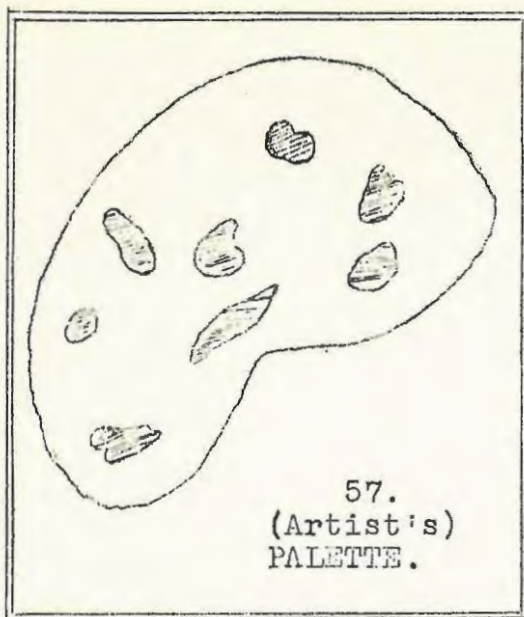
4481
"Church
window"
H2



7537 "Church
-- Lent"
Ho2, Fo2, ----.



7421 "Bishop's robe"
DFo3, Fox2, Fo2.



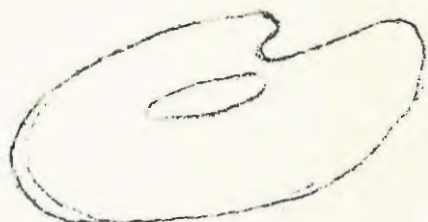
57.
(Artist's)
PALETTE.



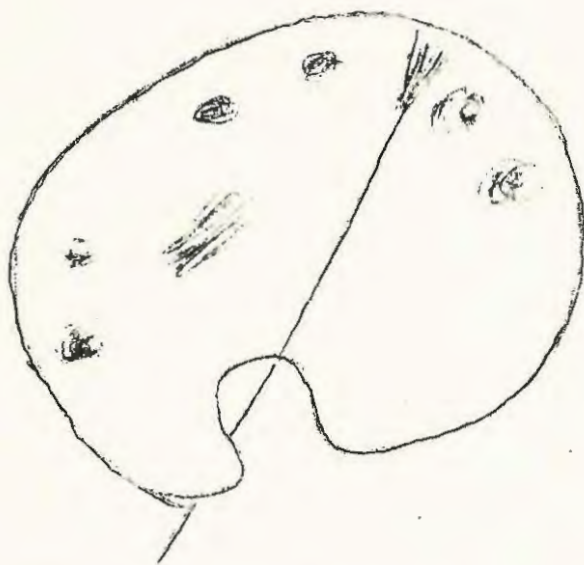
19325 "Pallette"
AF3, AF3, AFo3.



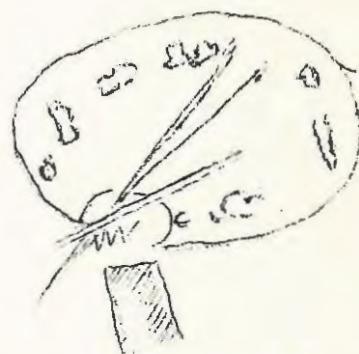
13725 "Palette with
Brushes." AFo3, AFo3
AFo3.



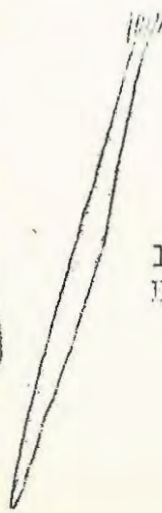
18131. "Pattern" (?)
AF3, AF2, AFo3.



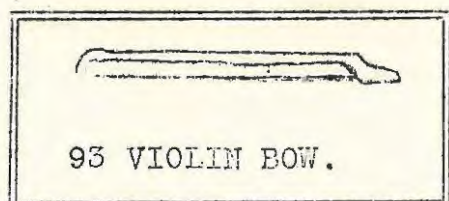
9064 "Palette and Paint
Brush". AFo3, AF3, AFo3.



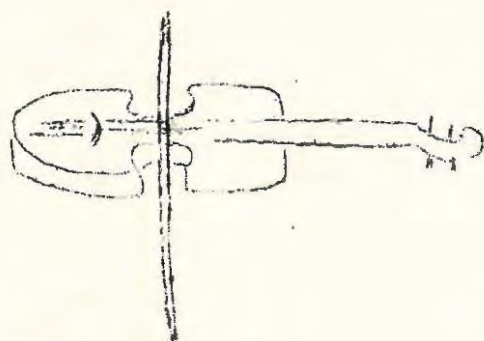
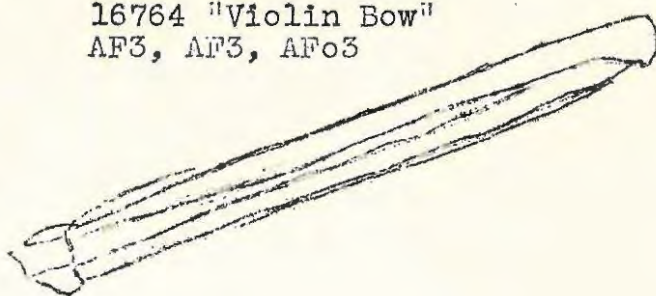
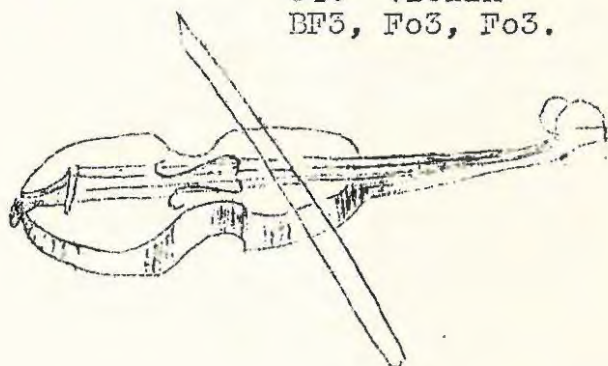
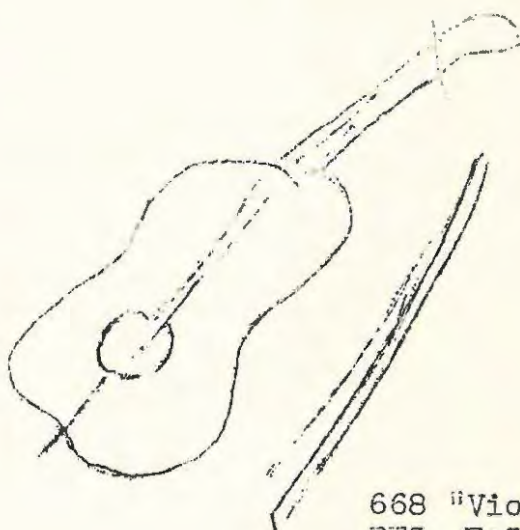
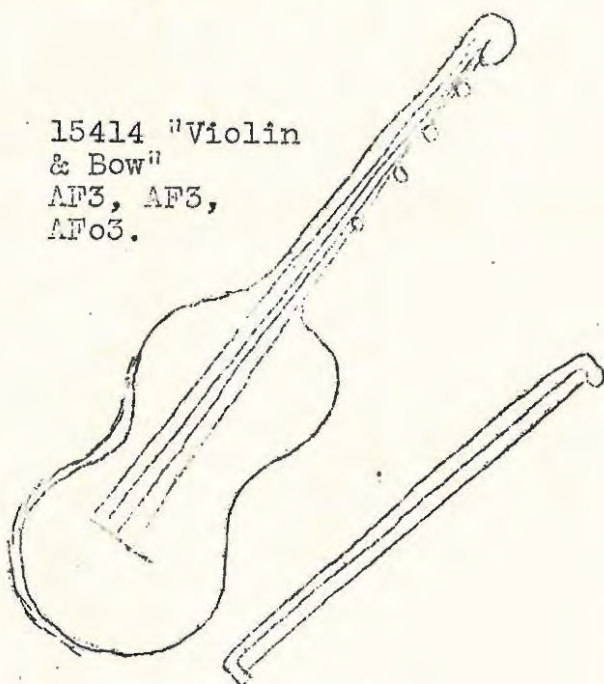
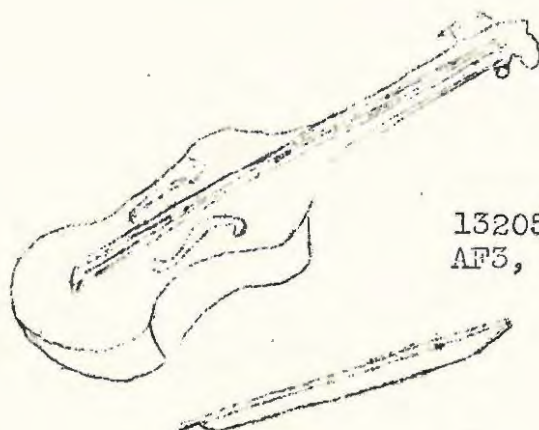
11256 "Palette &
Paints".
AFo3, AF3, AFo3.

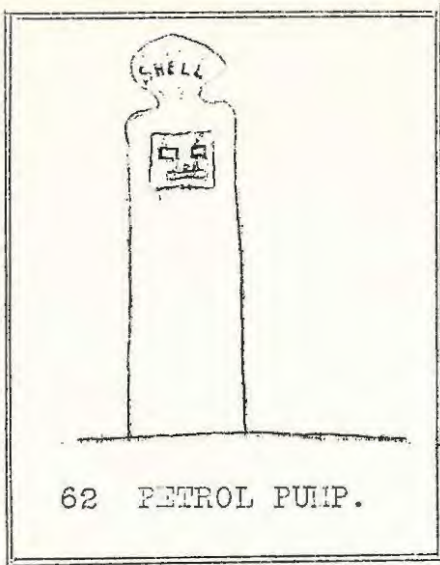


14665. "Palette and
Brush. AFo3, AF3, AFo3



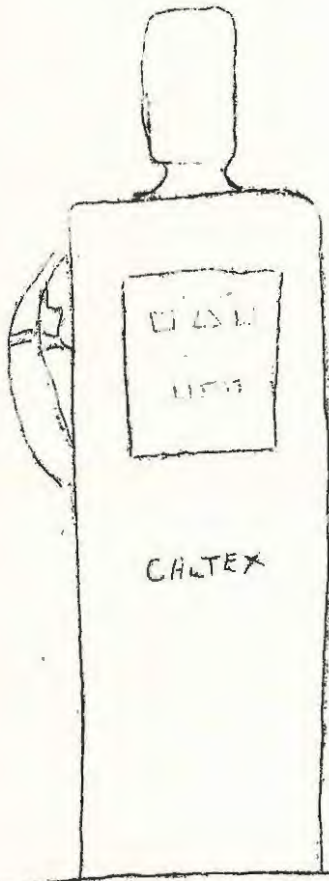
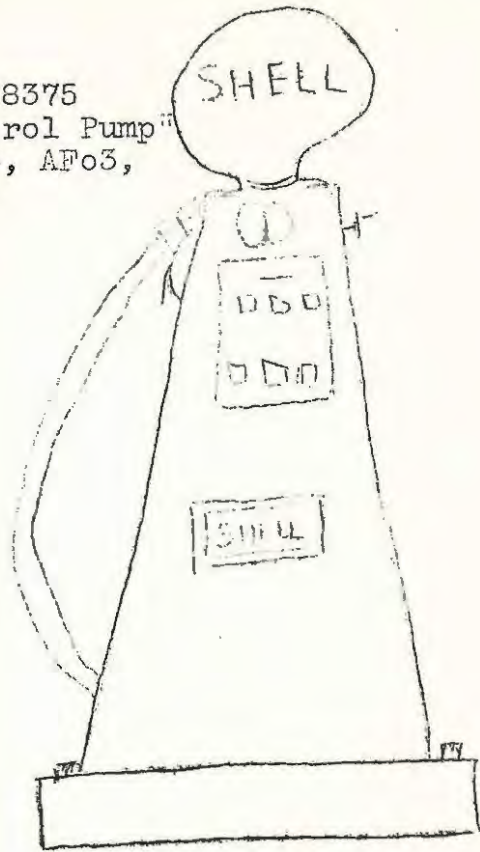
93 VIOLIN BOW.

16764 "Violin Bow"
AF3, AF3, AFo33316 "Violin"
BF3, Fo3, Fo3.540 "Violin"
BF3, Fo3, Fo3.15414 "Violin
& Bow"
AF3, AF3,
AFo3.668 "Violin"
BF3, Fo3,
Fo3.13205 "Violin & Bow"
AF3, AFo3, AFo3.



62 PETROL PUMP.

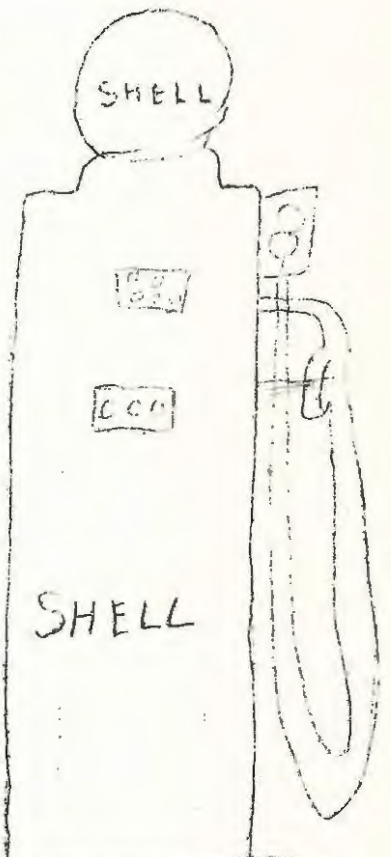
8375
"Petrol Pump"
AFo3, AFo3,
AF3.



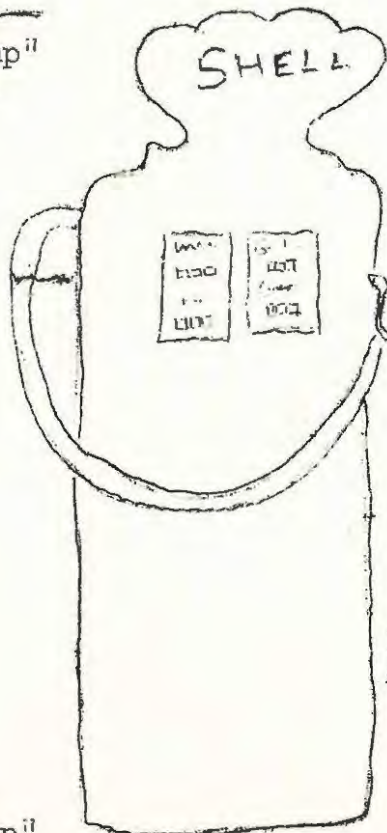
1323 "Petrol Pump"
AFo3, AFo3, AF3.



4323
"Petrol Pump"
AF3, AFo3,
AF3.



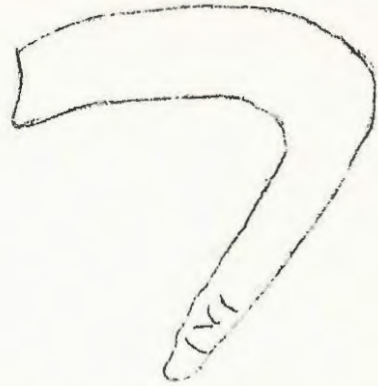
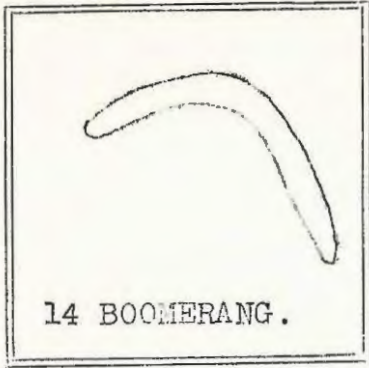
2902 "Petrol Pump"
AFo3, AFo3, AF3.



15338, "Petrol Pump"
AFo3, AFo3, AF3.



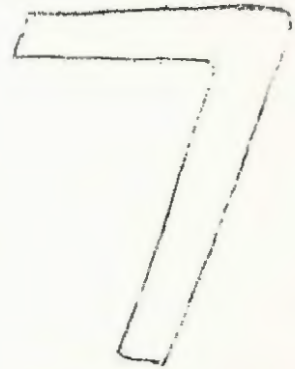
3689 "Petrol Pump"
AFo3, AFo3, AF3.



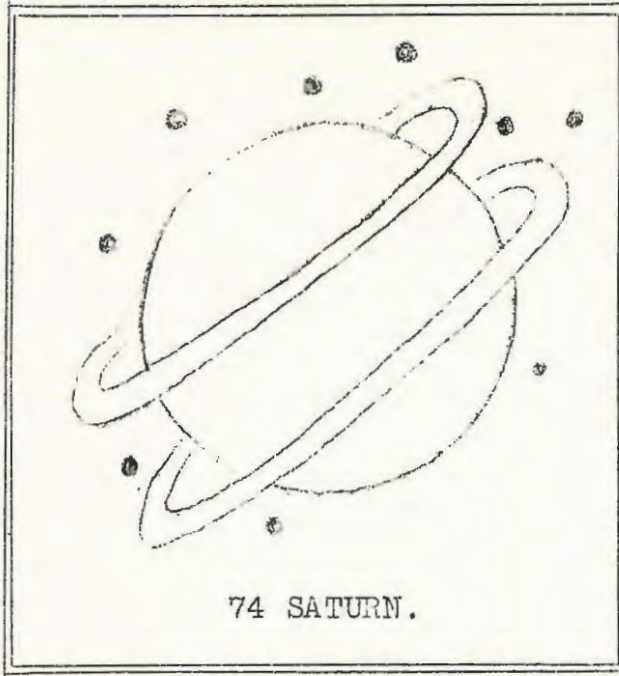
5650 "Boomerang"
AF3, AF3, AF3.



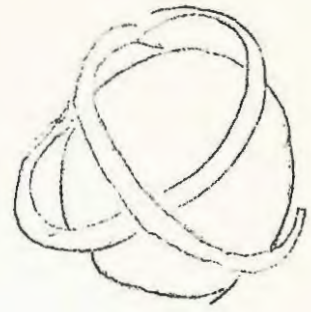
8709 "Boemerang"
AF3, AF3, AF3.



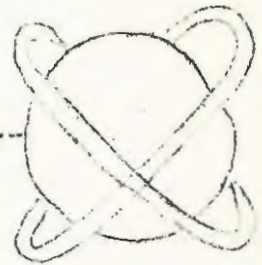
18606 "Figure Seven"
Fx2, Hx1, F3.



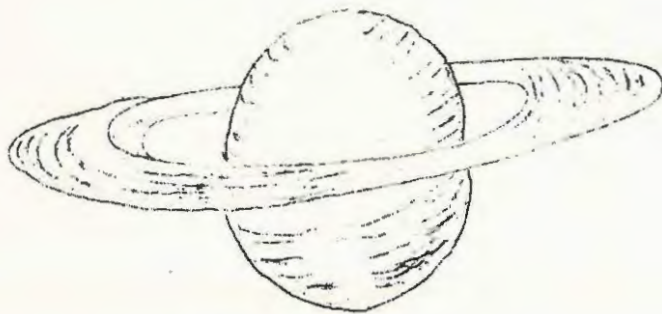
74 SATURN.



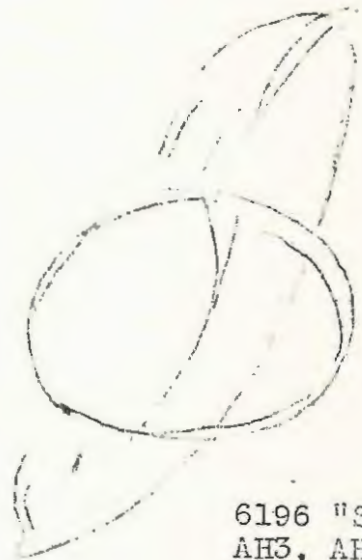
Better effort



15716 "Planet ?"
EFx3, BFx3 BHo3.



11512 "Saturn"
AH3, AH3, AH3.



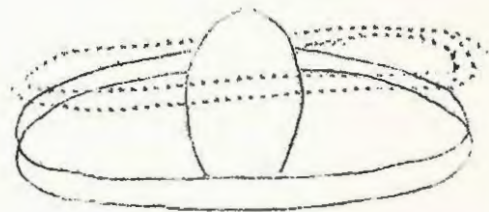
6196 "Saturn"
AH3, AH3, AH3.



10420 "Saturn"
AH3, AH3, AH3



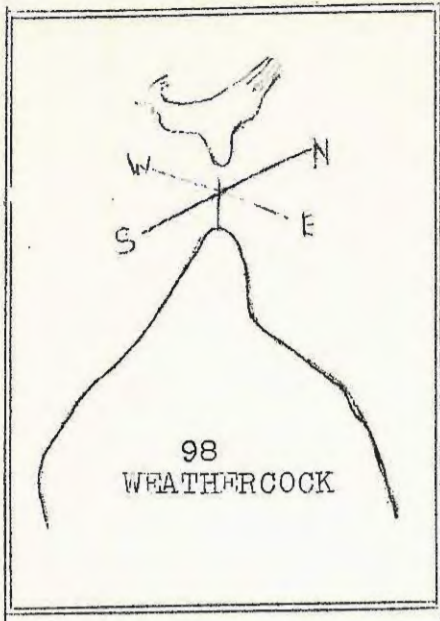
4443 "Moon"
H3J, H3J, H3.



15645 "Planet"
EHLx3, BH2, BH3.



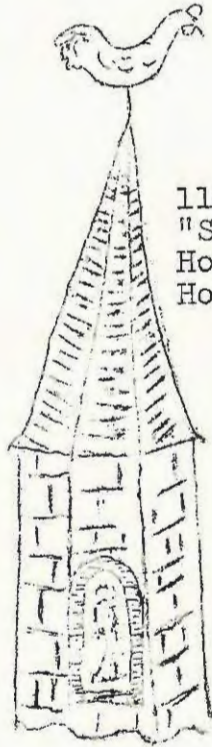
9658 "World"
BH3, H3, CH3.



98
WEATHERCOCK



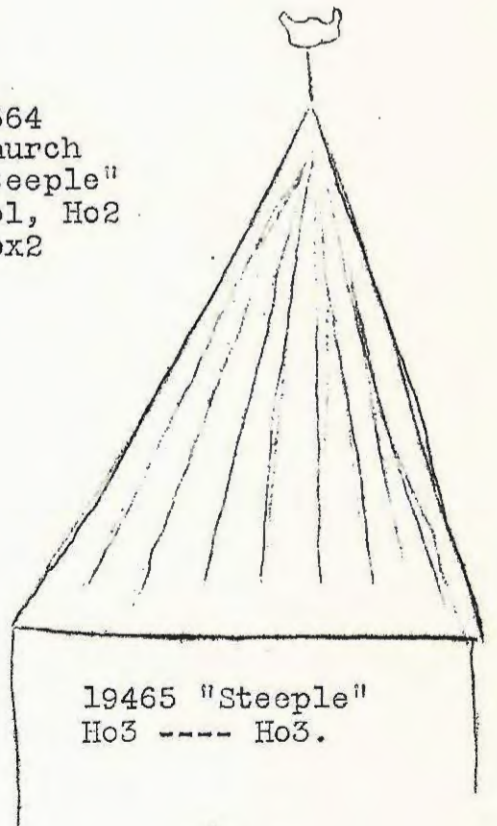
17806
"Weather Cock"
AH3, AF3,
AF3.



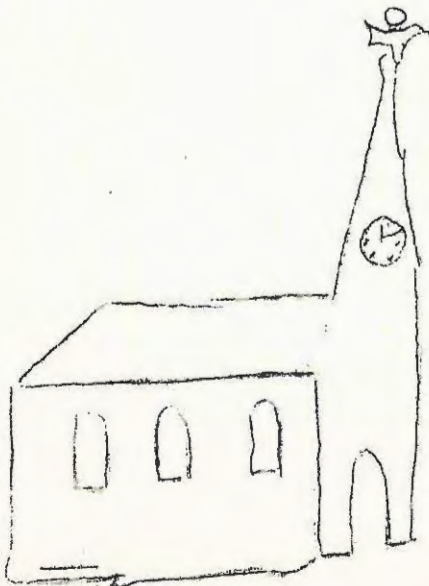
11025
"Steeple"
Ho3, Fo3,
Ho3.



18664
"Church
Steeple"
Ho1, Ho2
Hox2



19465 "Steeple"
Ho3 ---- Ho3.

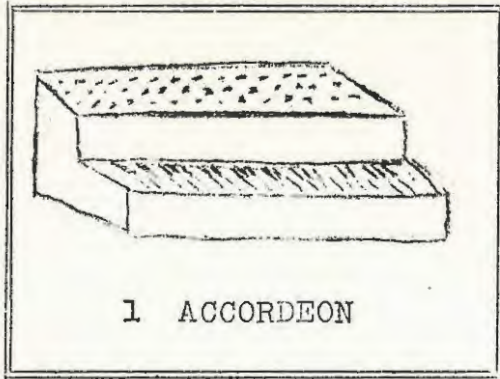


16419.
"Church & Clock"
Ho2, Ho2, Hox2.

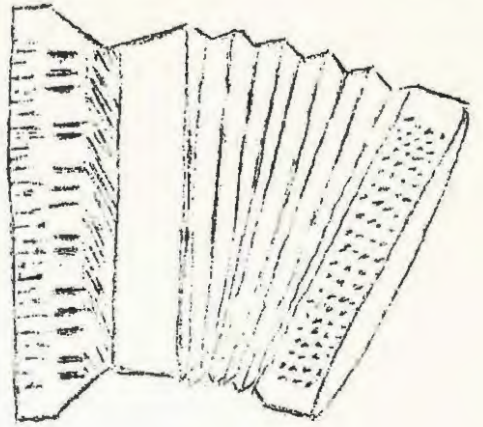


17467
"Church
Tower"
Ho2
Ho2
Hox2

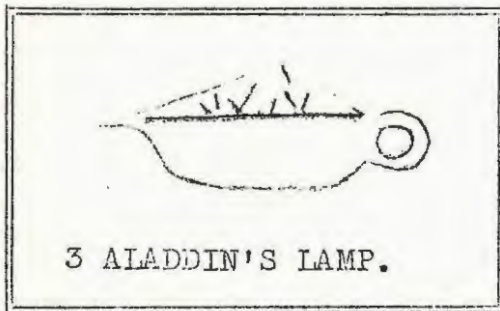
(6) HITS ON CONTROL SERIES
ORIGINALS



1 ACCORDEON



6753 "Accordion"
AFo3, AF3, AFx3.



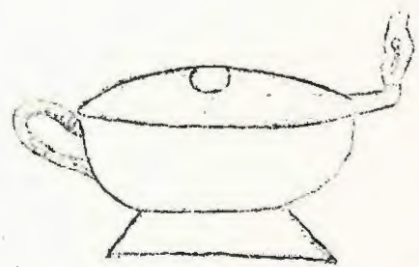
3 ALADDIN'S LAMP.



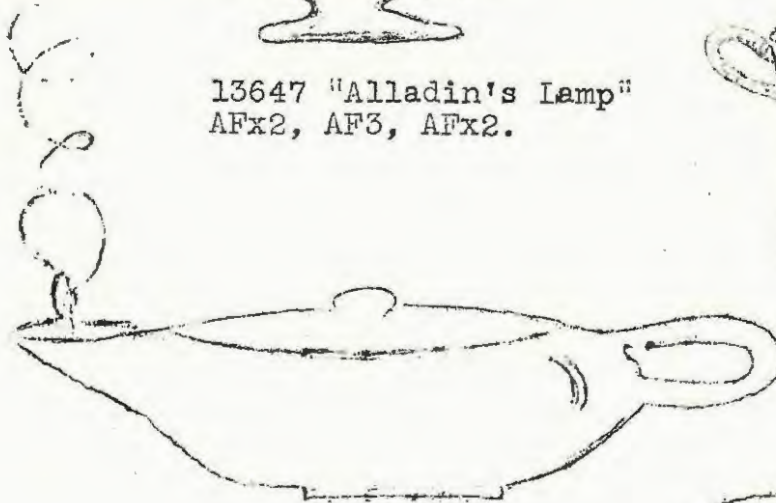
3166 "Aladdin's Lamp or
Pot-container."
AFx2, AF3. AFx2.



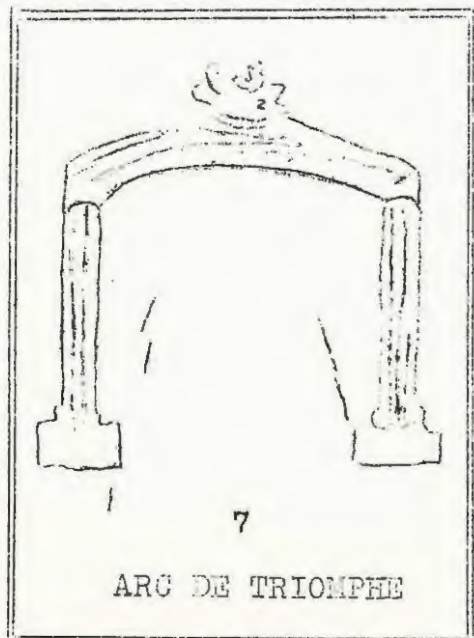
13647 "Alladin's Lamp"
AFx2, AF3, AFx2.



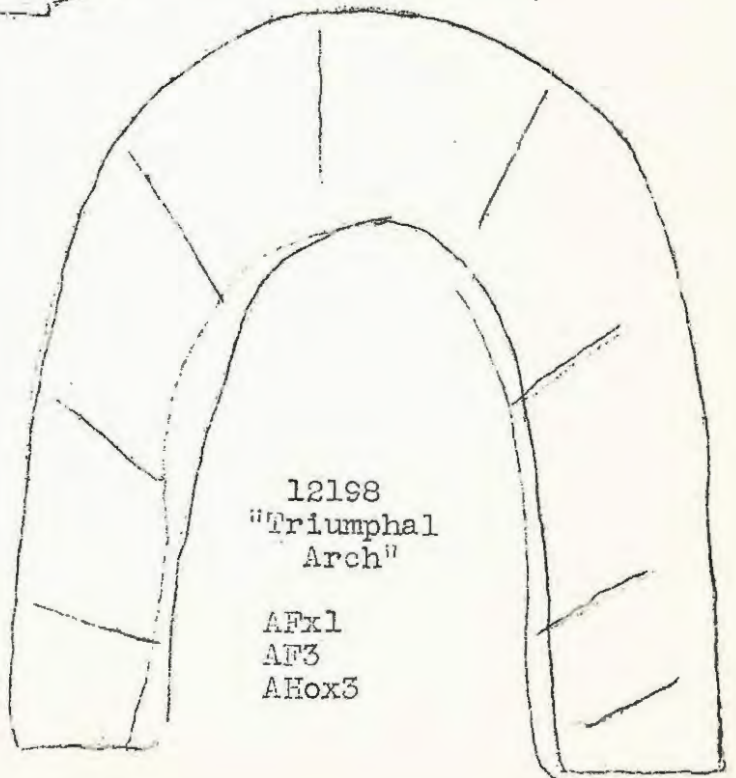
14075 "Alladin's
Lamp" AFo3, AF3, AF3



9823. "Aladdin's
Lamp" AF3, AF3, AF3.

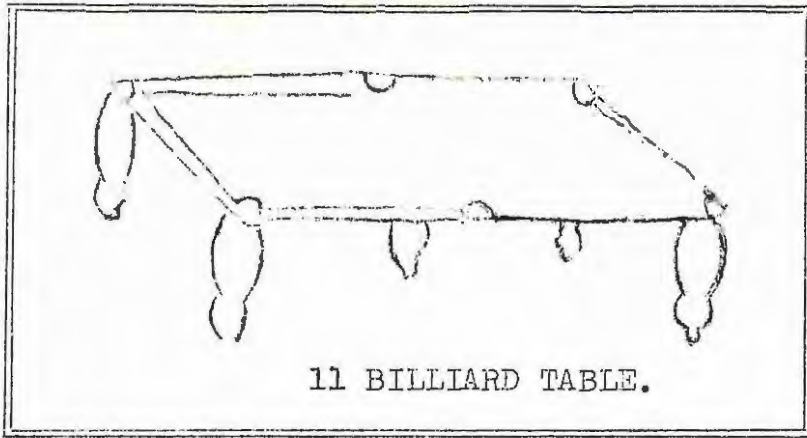


7
ARC DE TRIOMPHE

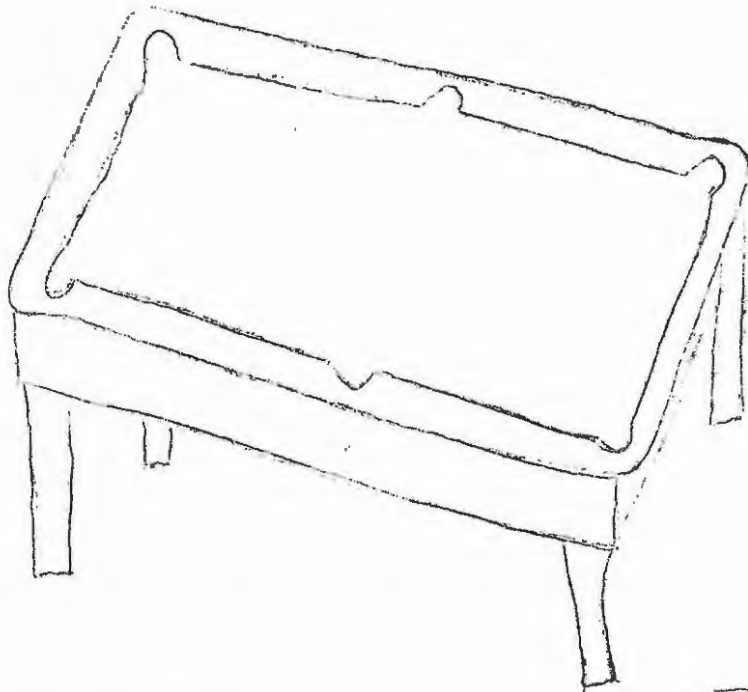


12198
"Triumphal
Arch"

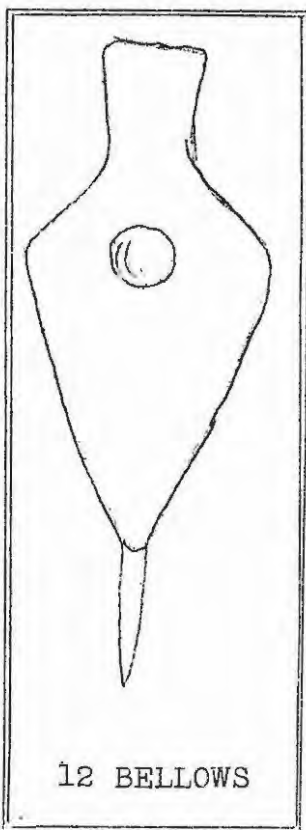
AFx1
AF3
AHox3



11 BILLIARD TABLE.

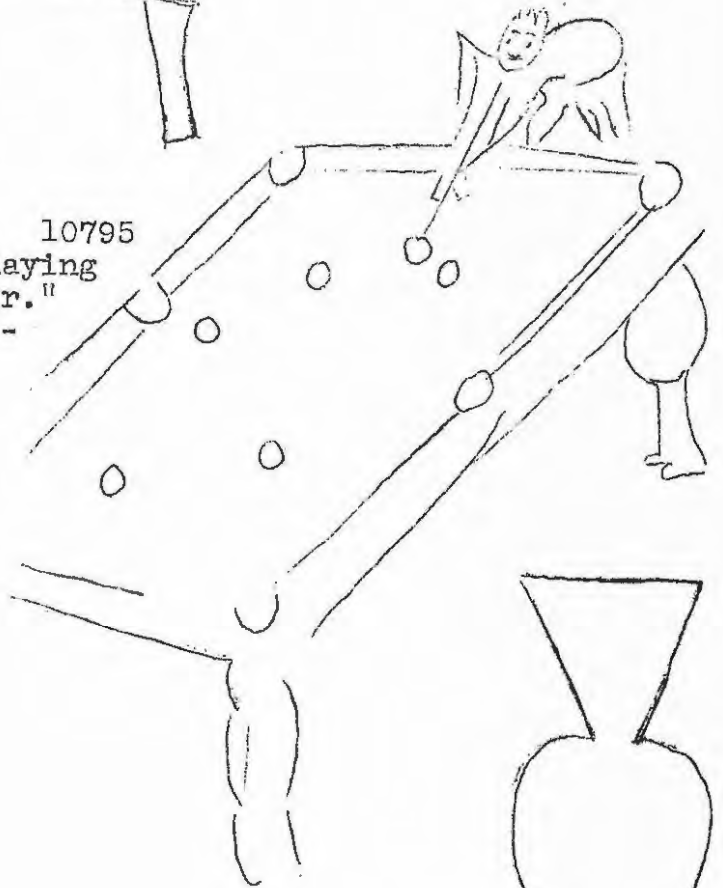


6240 "Billiard Table"
AF3, AF3, AF3.

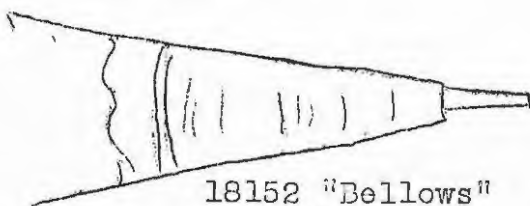
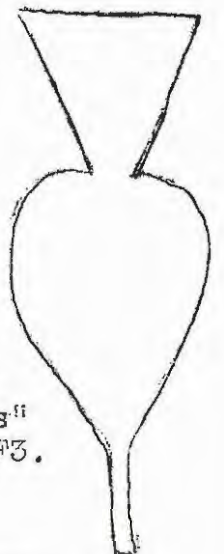


12 BELLOWS

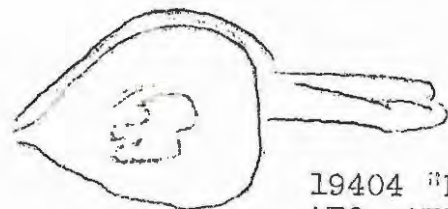
10795
"Man playing
Snooker."
Fx3 ---
Fo3.



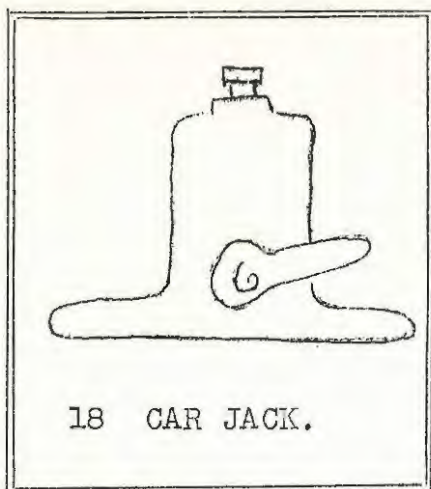
12070
"Bellows"
AF3, AF3, AF3.



18152 "Bellows"
AH2, AF2, AF2.

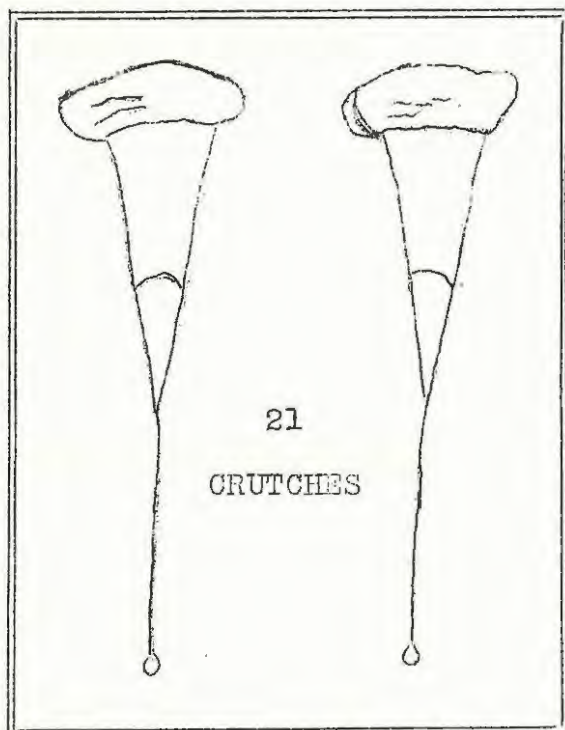
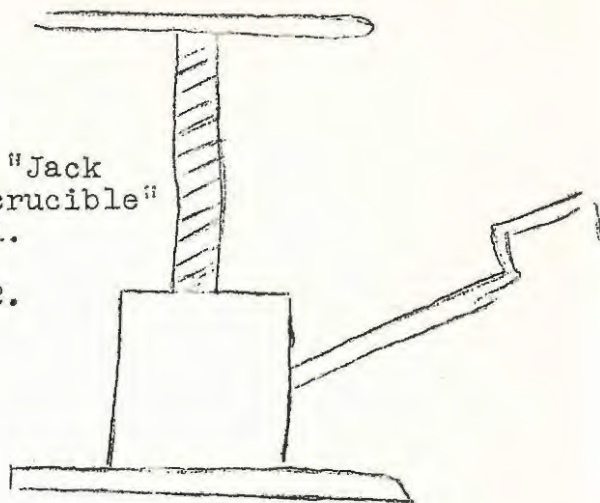


19404 "Bellows"
AF2, AF3, AF3.

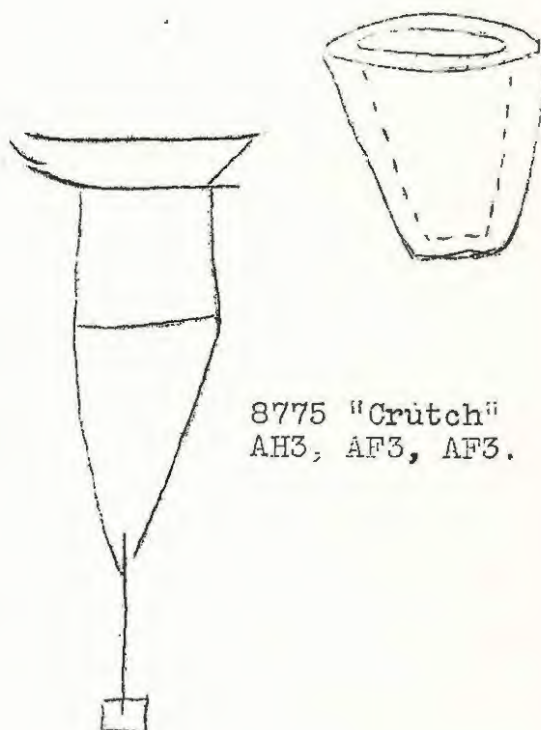


18 CAR JACK.

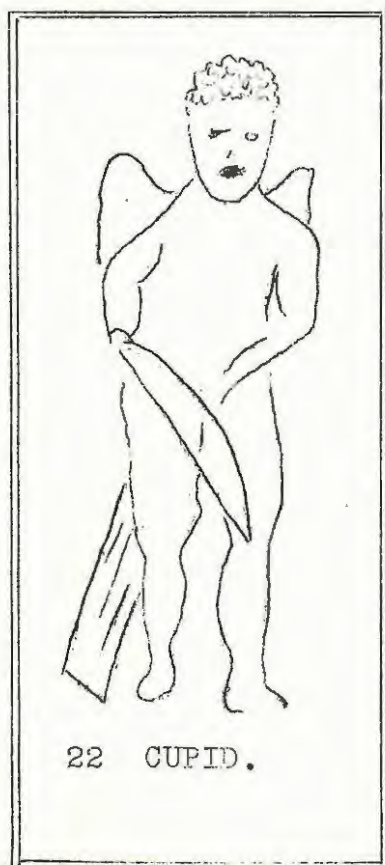
3630 "Jack and crucible"
AFox1.
AF3,
BFox2.



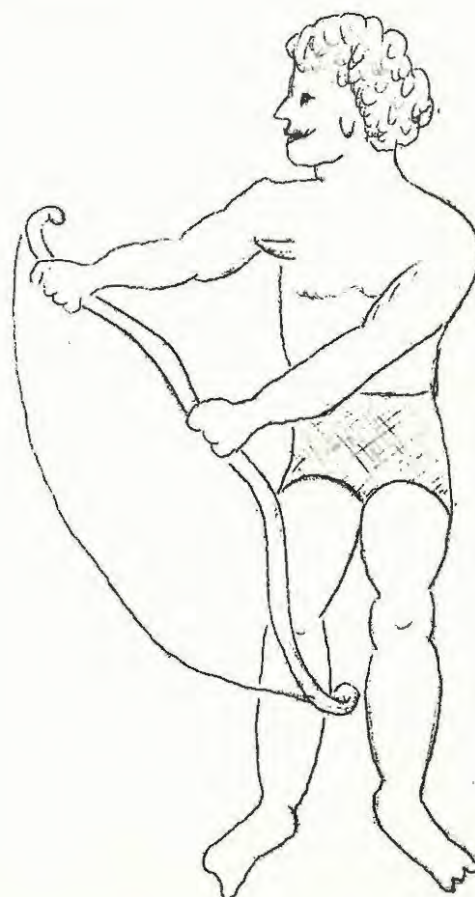
21
CRUTCHES



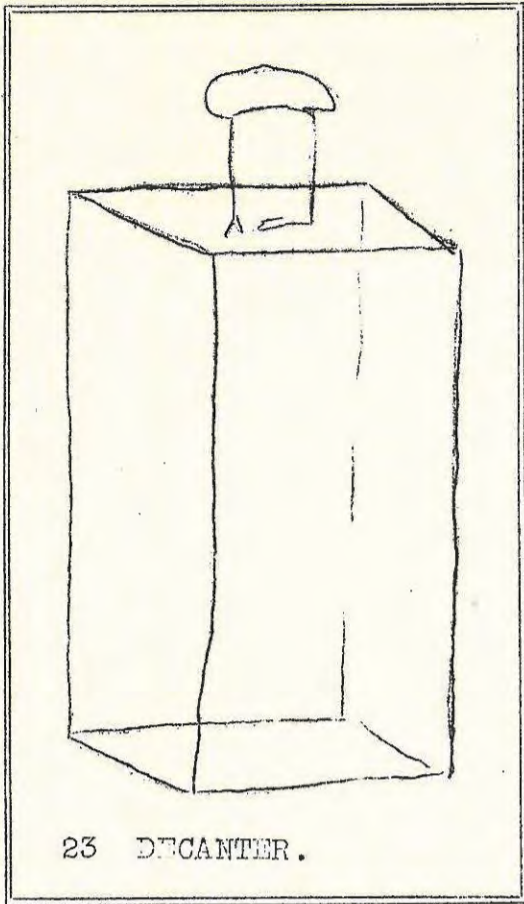
8775 "Crutch"
AH3, AF3, AF3.



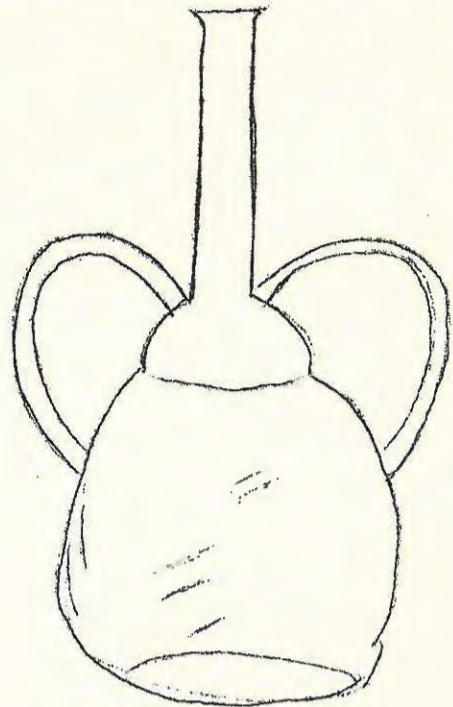
22 CUPID.



13399
"Cupid"
AH3,
AF2,
AH3.



23 DECANTER.



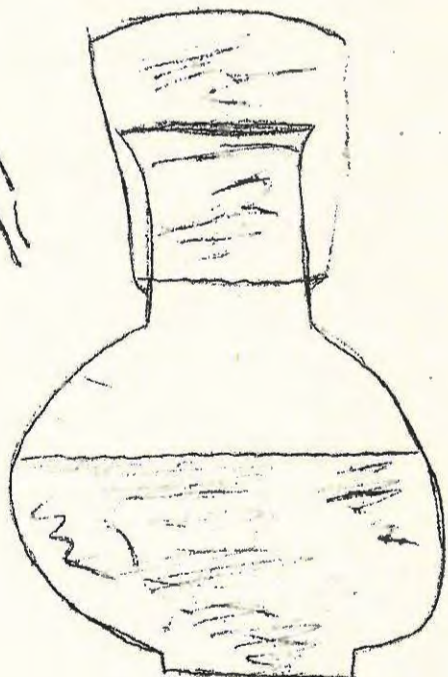
1061 "Decanter"
AFox1, ---, AF1



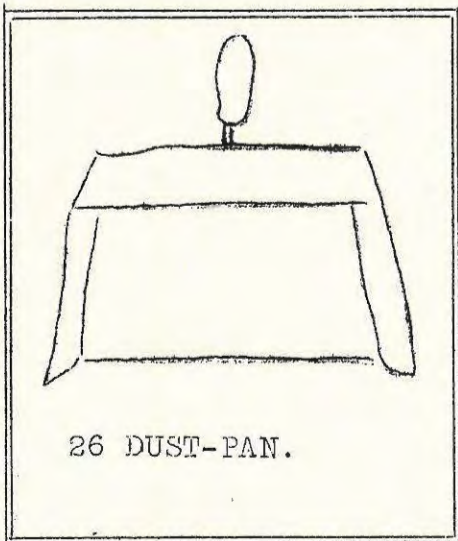
24 DISCUS
THROWER
(MODEL)



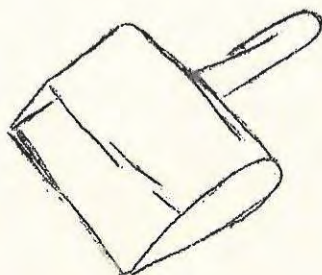
15455 "Discus
Hurler see MYRON"
AF3, AF3, AF3



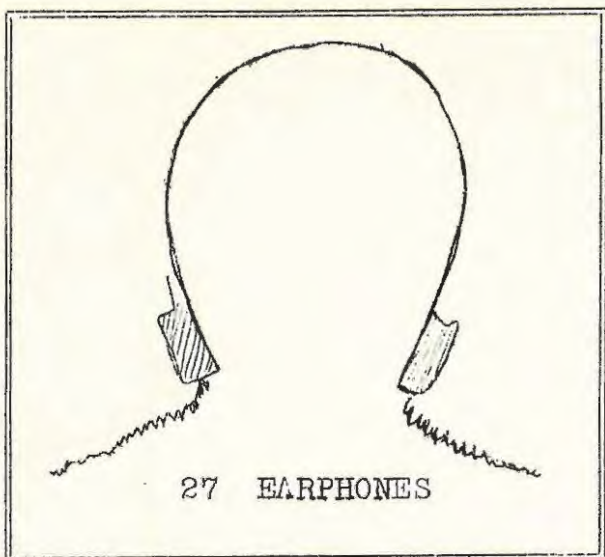
2114 "Decanter & glass"
AFox1, AFo2, AFox2.



26 DUST-PAN.



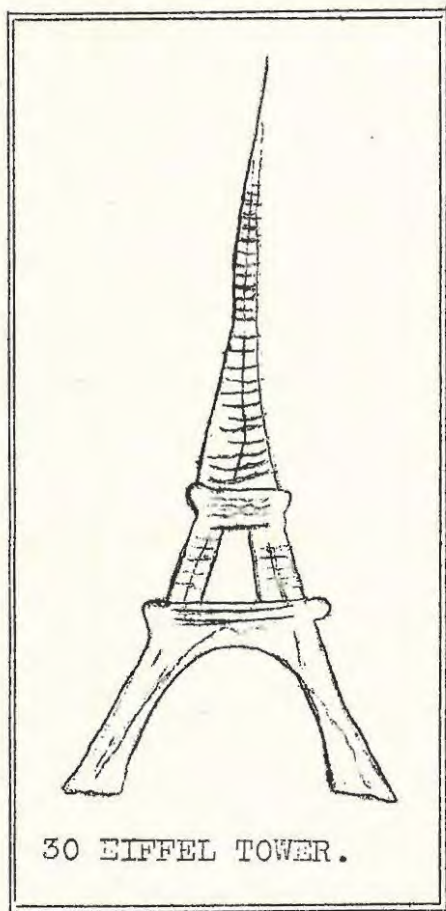
10720 "Dust pan"
AFo3, AF3, AF3.



27 EARPHONES

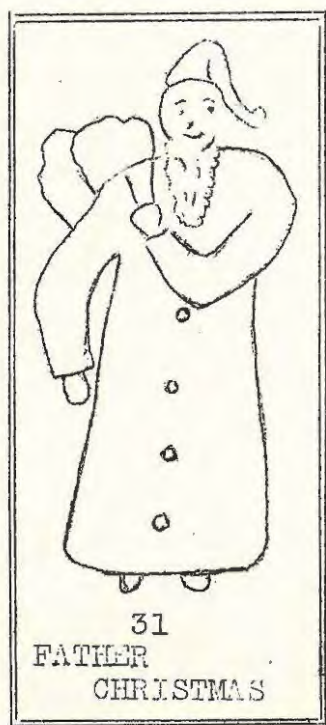
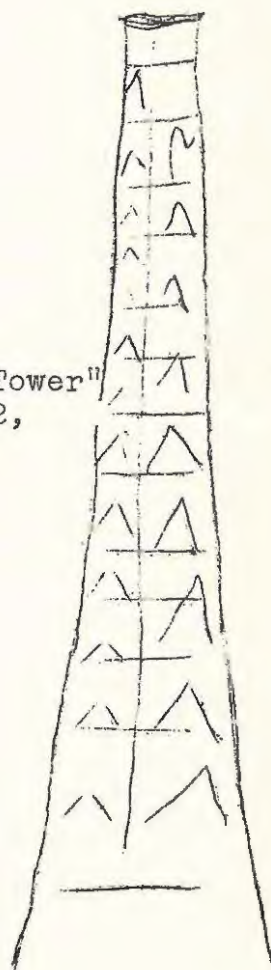


15720. "Earphones"
AFo2, AF2, AF3.



30 EIFFEL TOWER.

17672
"Eiffel Tower"
AFo2, AF2,
AF3.



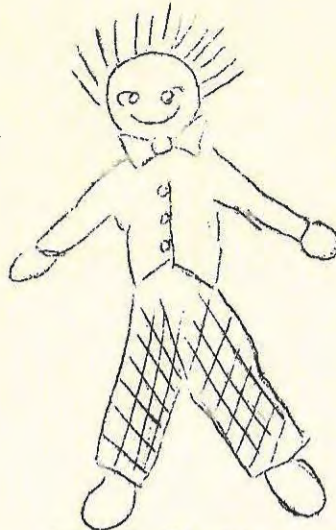
31
FATHER
CHRISTMAS



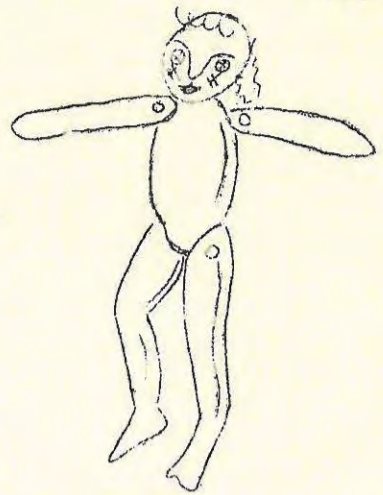
17805
"Father Christmas"
AH2, AF3, AFox2.



40 GOLLYWOG.



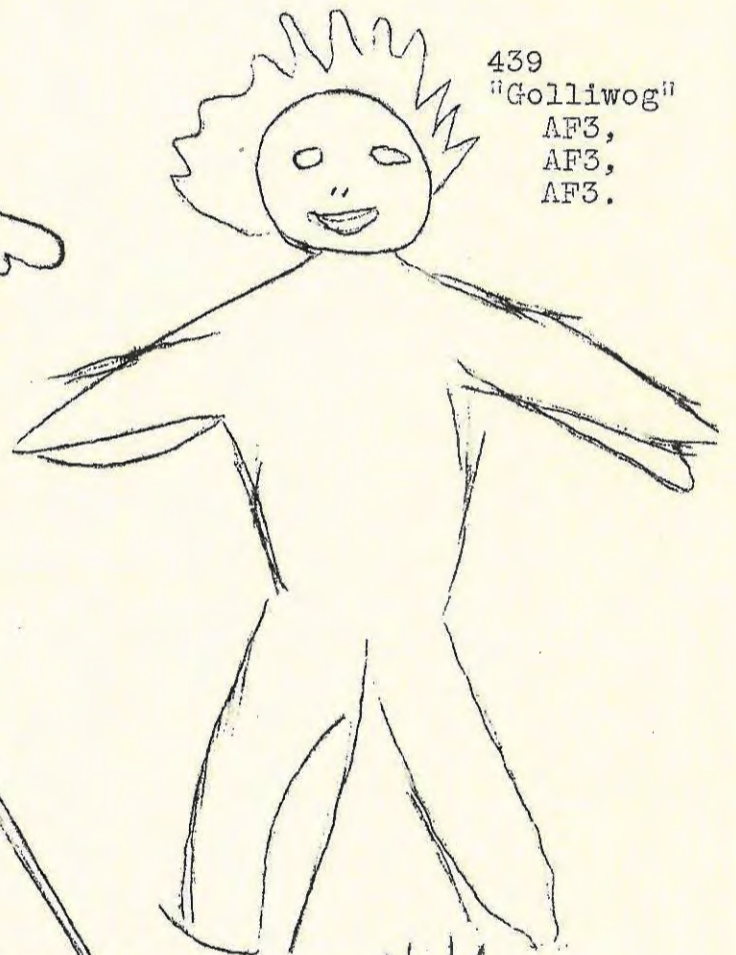
17721 "Gollywog"
AF3, AF3, AF3.



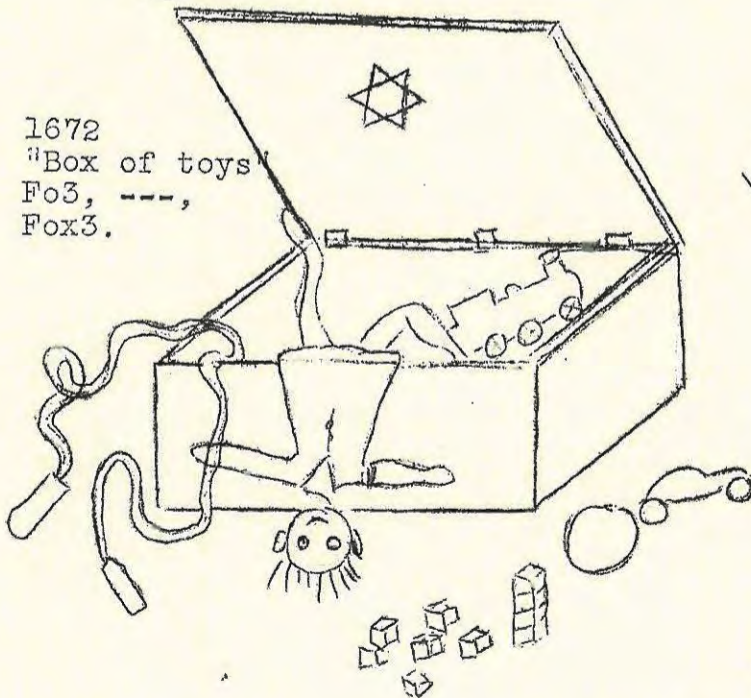
2134 "Doll"
J, ---, J.



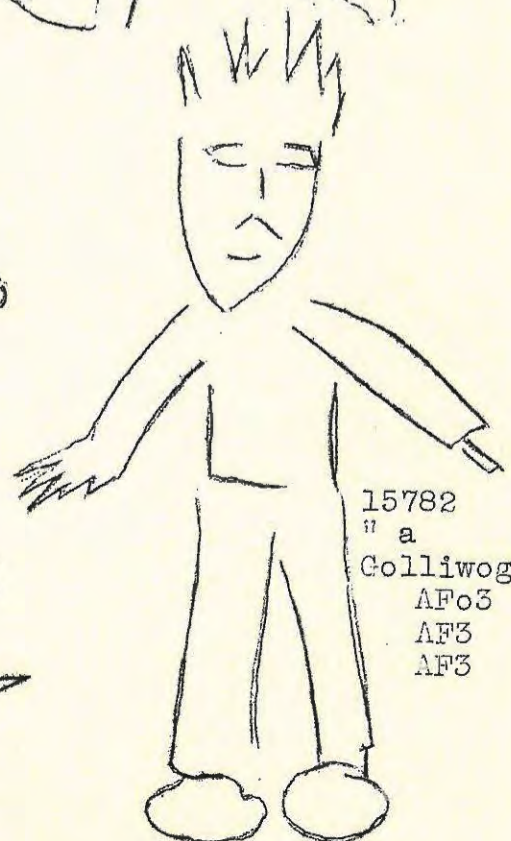
9578
"Golliwog"
AFo3,
AF3,
AF3.



439
"Golliwog"
AF3,
AF3,
AF3.



1672
"Box of toys"
Fo3, ---,
Fox3.



15782
" a
Golliwog"
AFo3
AF3
AF3

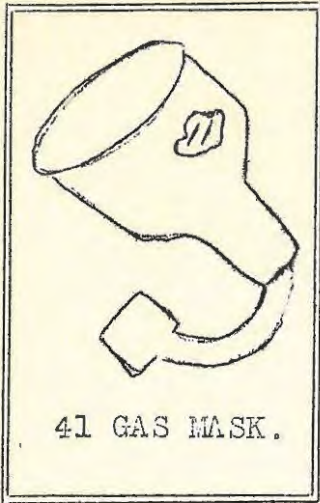
3004 "Football, bones
Fence Doll"



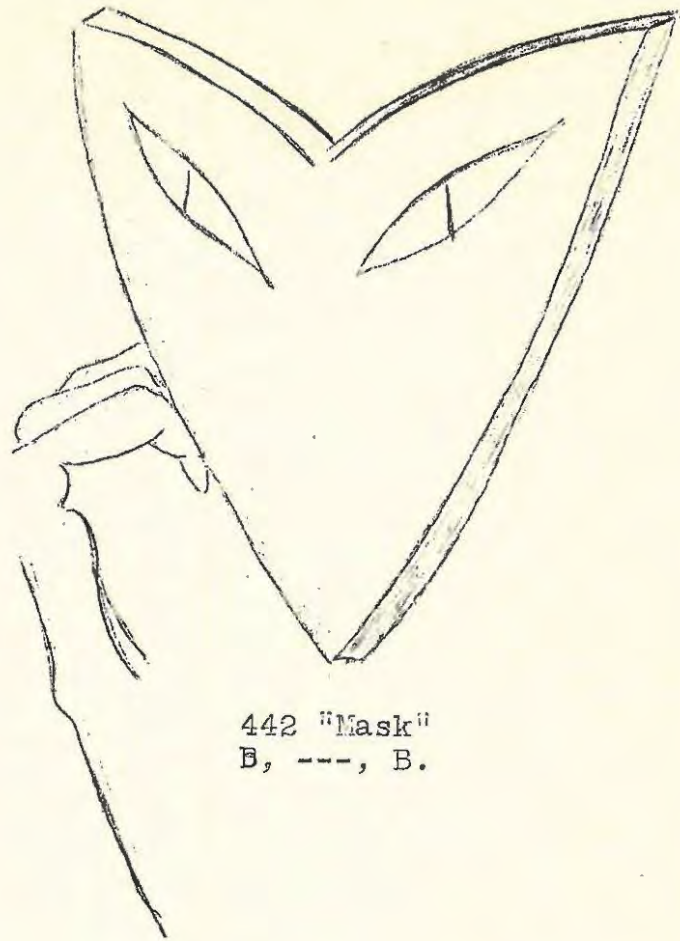
Fo3,
Fox2,



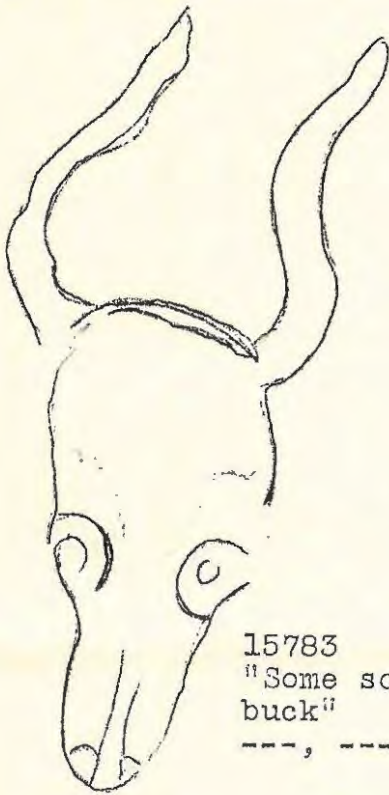
5269
"Golliwog"
AF3, AF3,
AF3.



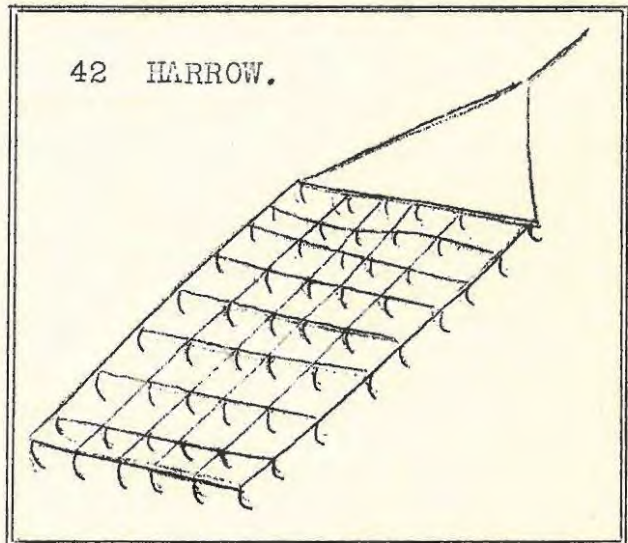
41 GAS MASK.



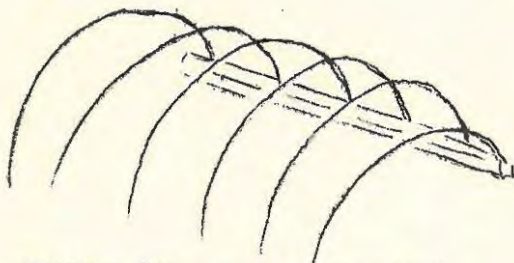
442 "Mask"
B, ---, B.



15783
"Some sort of
buck"
---, ---, Hox2

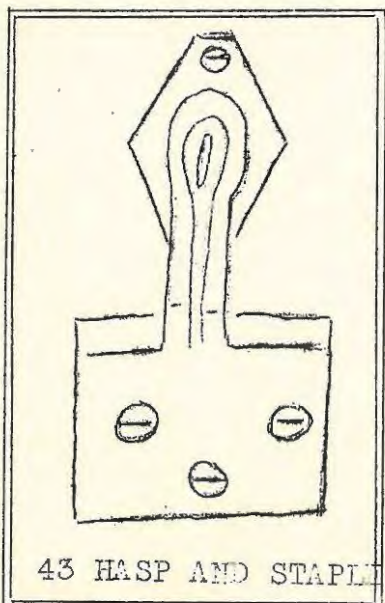
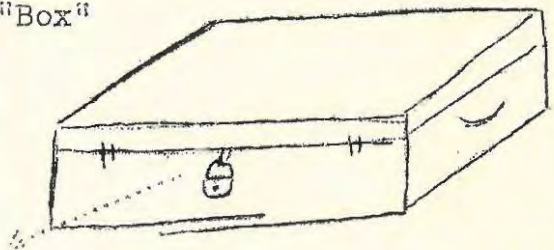


42 HARROW.

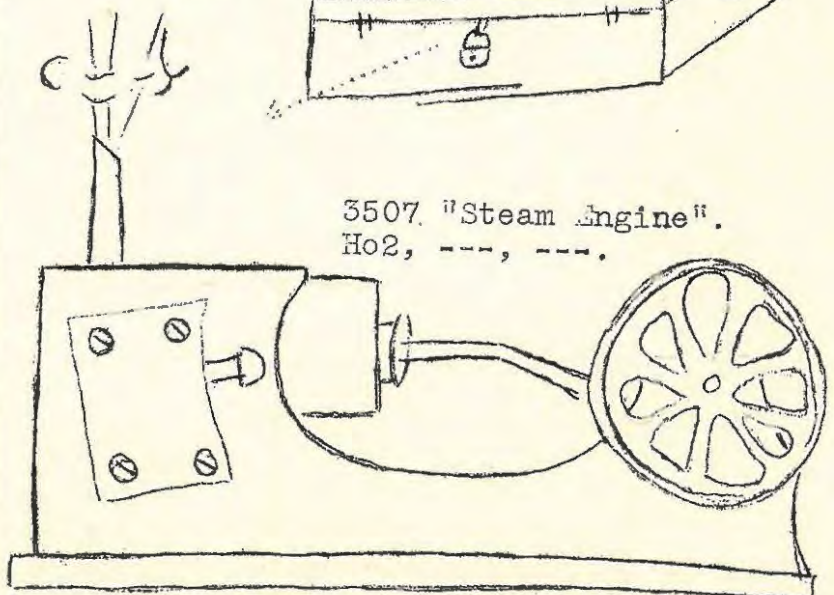


17030 "Harrows Teeth"
C, C, Hx3.

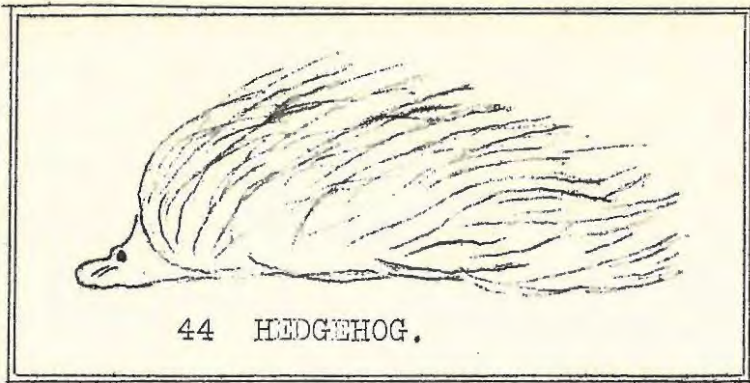
1749 "Box"
Hox1



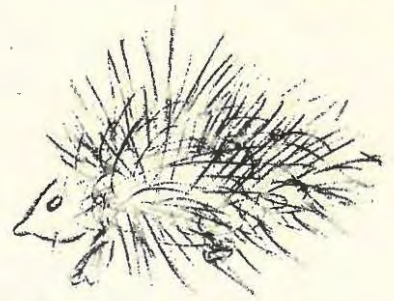
43 HASP AND STAPLE



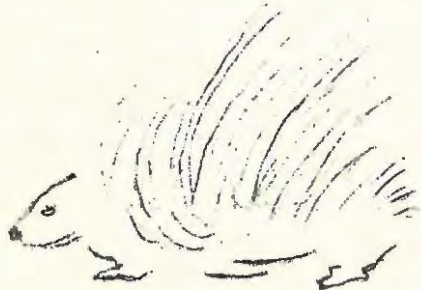
3507. "Steam Engine".
Ho2, ---, ---.



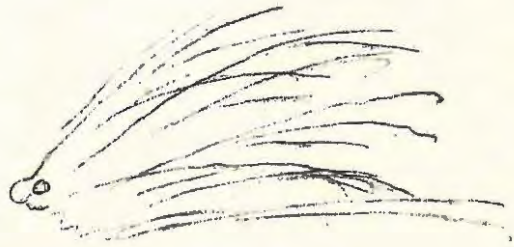
44 HEDGEHOG.



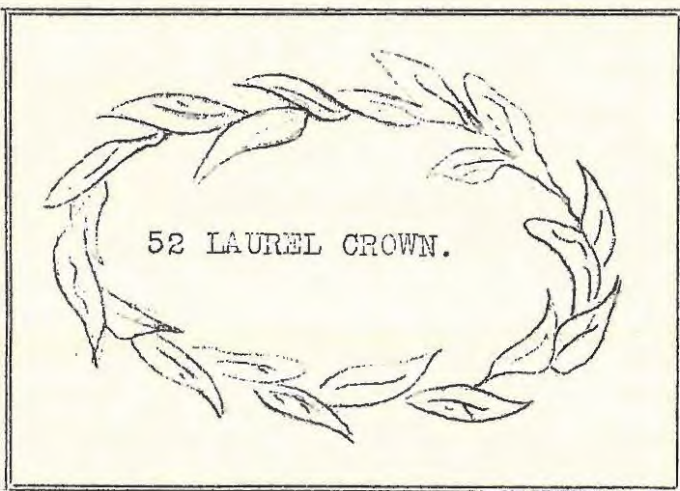
8152 "Porcupine"
EF3, EF3, EF3.



724 "Porcupine"
EF3, EF3, EF3.



13320 "Hedgehog"
AF3, AF3, AF3.



52 LAUREL CROWN.



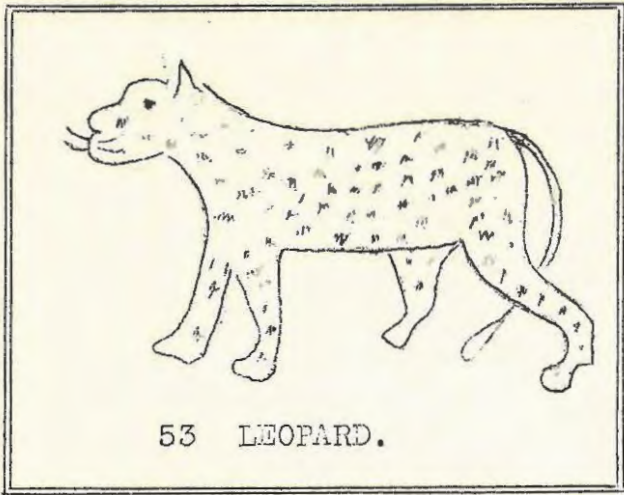
12806 "Porcupine"
EF3, EF3, EF3.



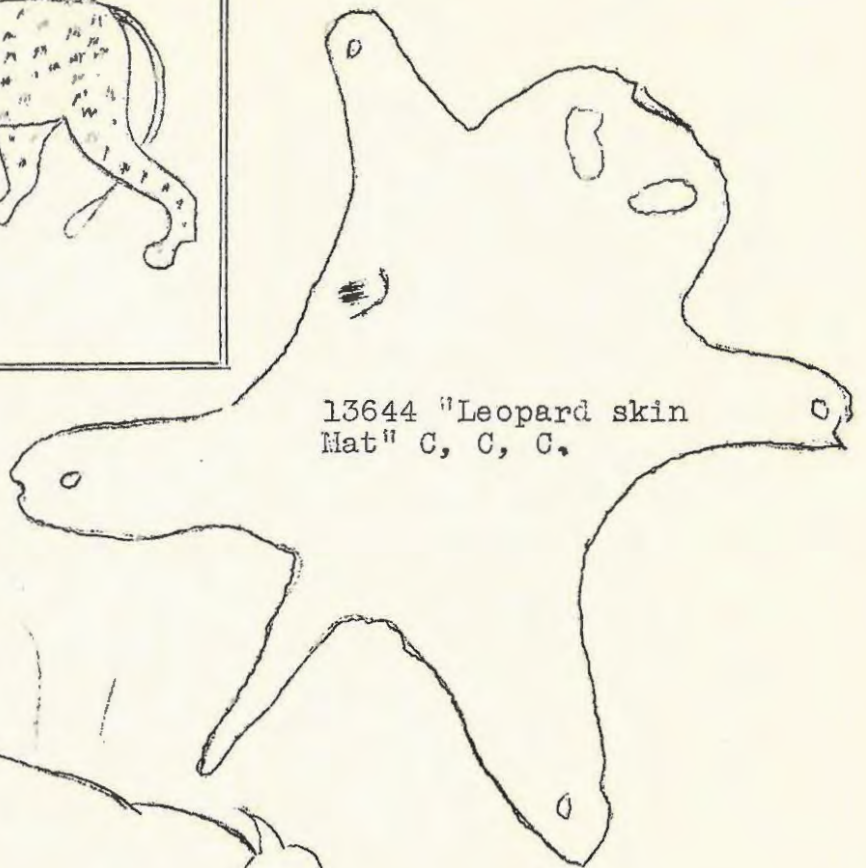
15413 "Laurel Wreath"
EF3, EF3, EF3.

14418 "Thistle
and Laurel Wreath"
AH2, ---, AFox2.

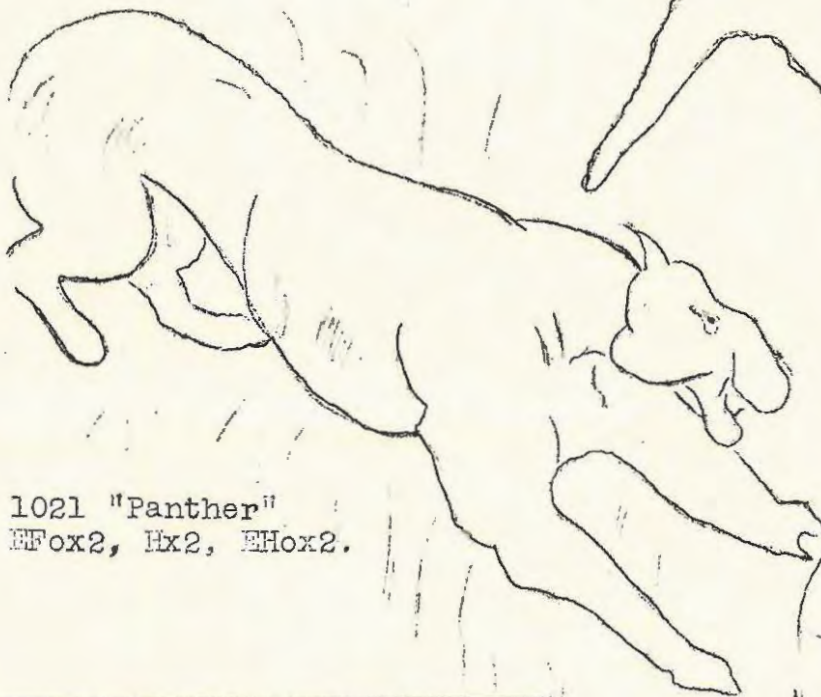




53 LEOPARD.



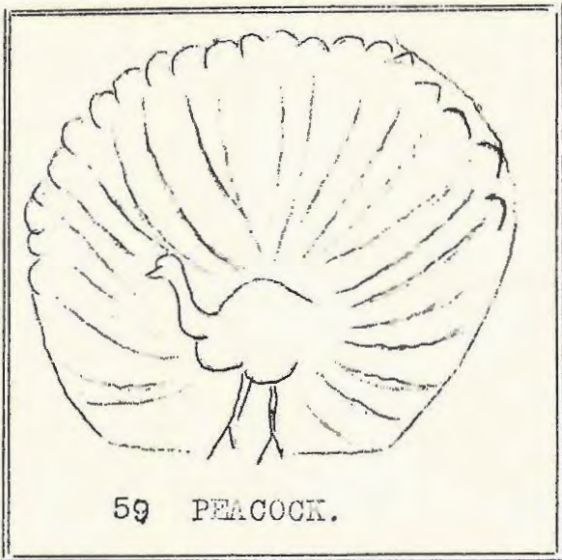
13644 "Leopard skin Mat" C, C, C.



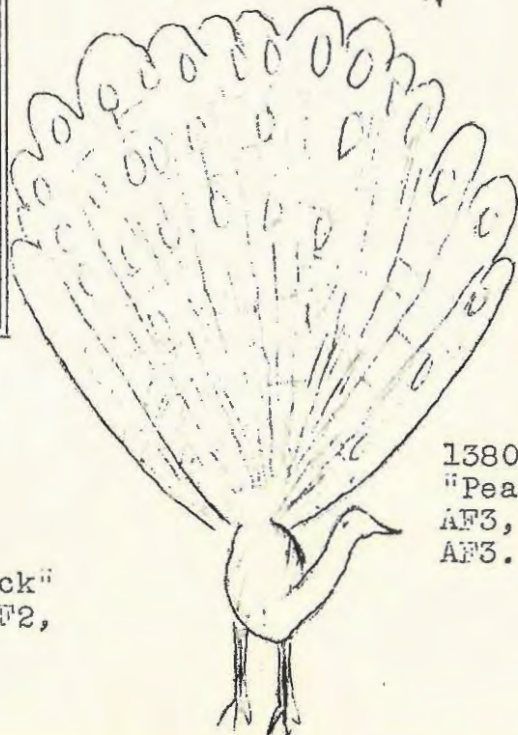
1021 "Panther"
EFox2, Hx2, EHox2.



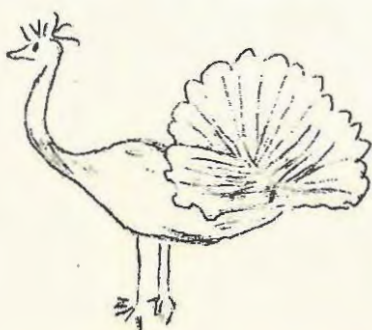
17319
"Peacock"
AF3, AF3, AF3.



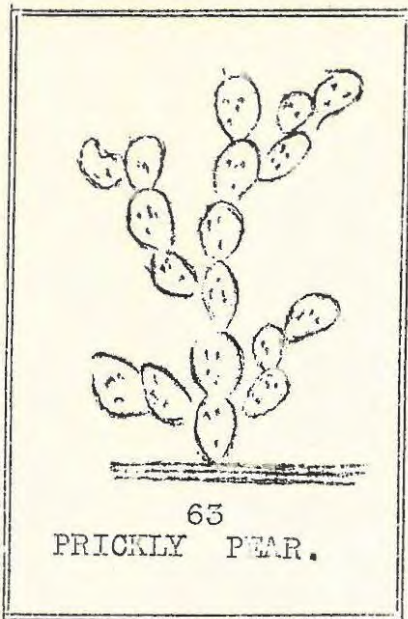
59 PEACOCK.



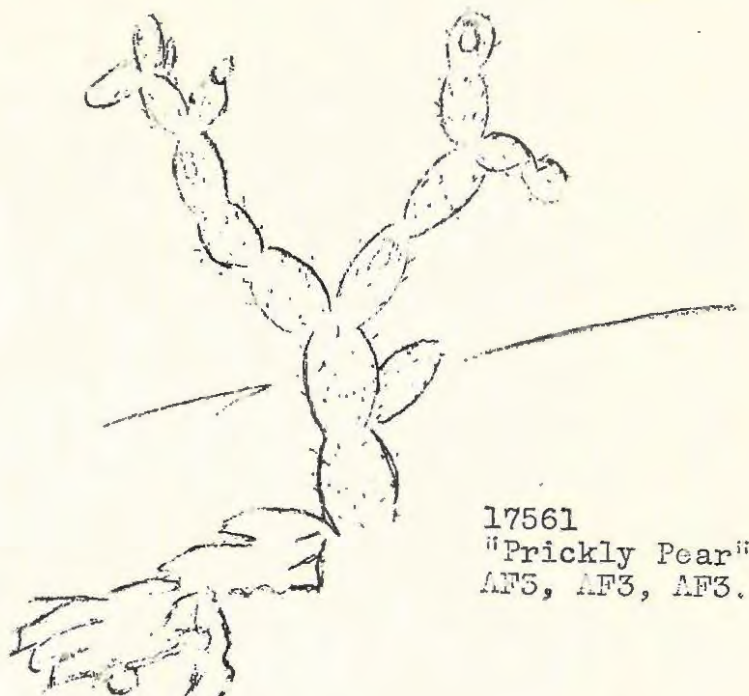
13806
"Peacock"
AF3, AF3,
AF3.



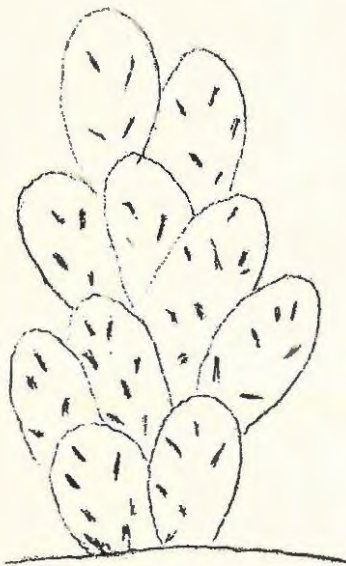
4318
"Peacock"
AF2, AF2,
AF3.



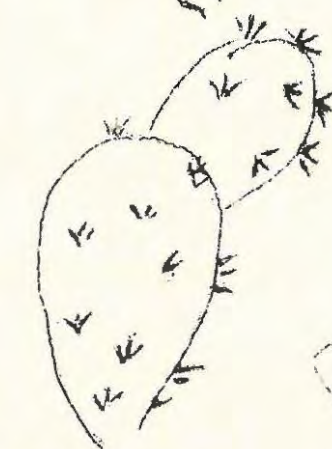
65
PRICKLY PEAR.



17561
"Prickly Pear"
AF3, AF3, AF3.



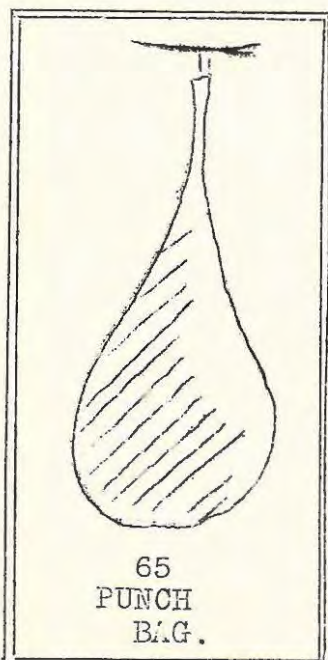
11384 "Prickly pear
Plant". AFox2, AF3,
AF3.



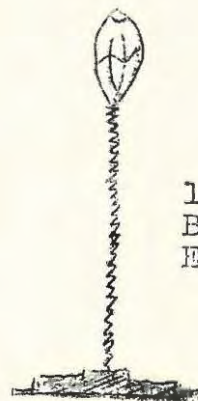
19569 "Prickly
Pear. AH3, AH3,
AF3.



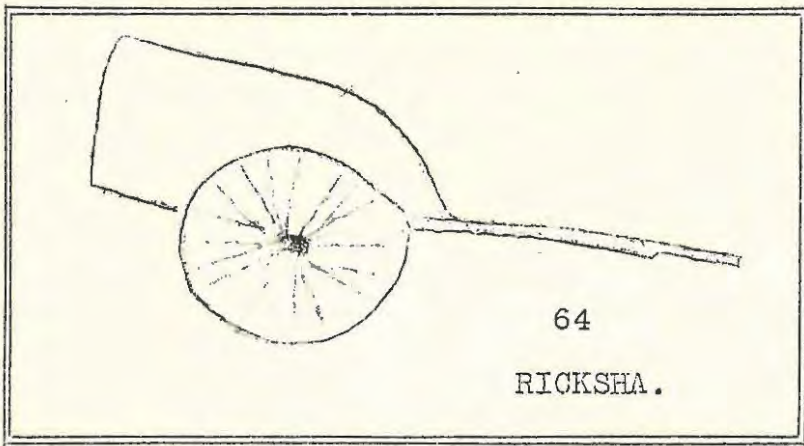
2095 "Prickly Pear"
AH3, AH3, AF3.



65
PUNCH
B.A.G.

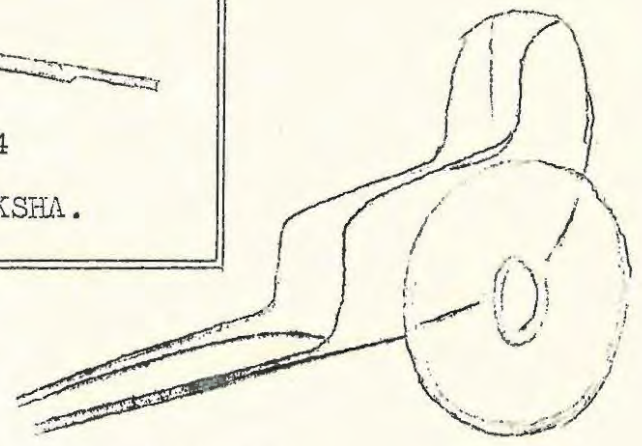


18310 "Punch
Ball".
EFx2, Fo3, EFx2.



64

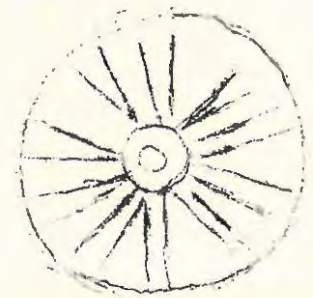
RICKSHA.



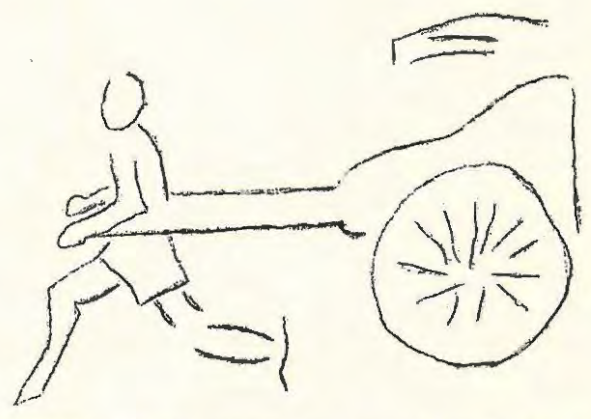
1461 "Rickshaw"
AFo3, AF3, AF3.



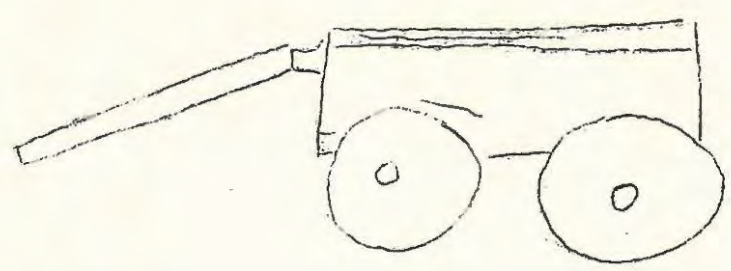
3258 "Toy cart"
Fx3, Fx1, BFo2.



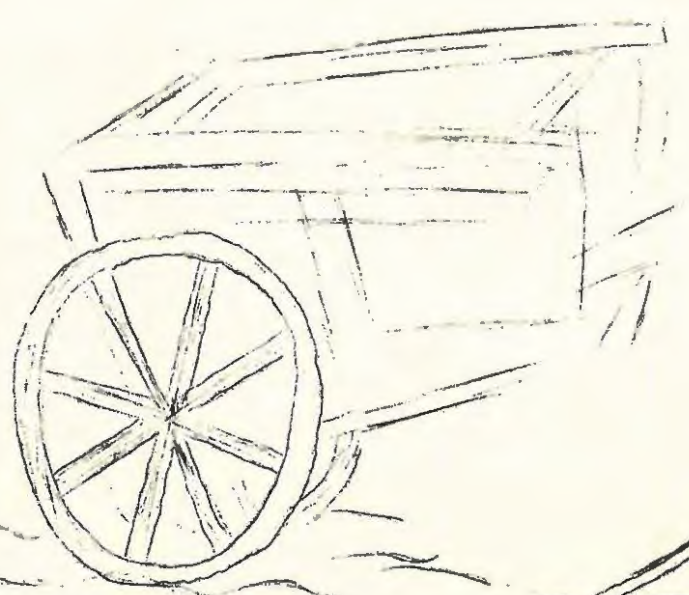
1058
"Wagon Wheel"
Ho1, H2, ----.



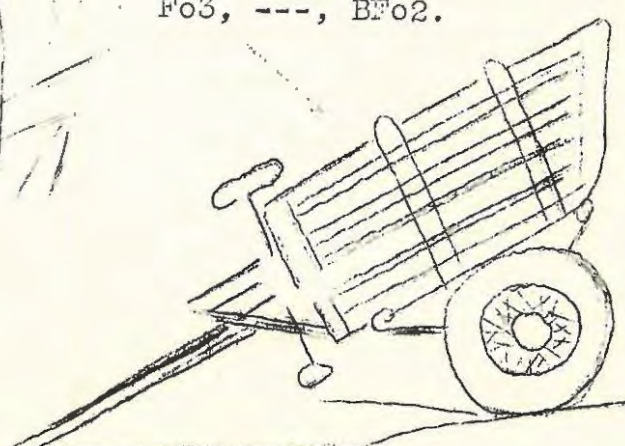
3547 "Man drawing a rickshaw"
AFo3, AFo3, AFo3.



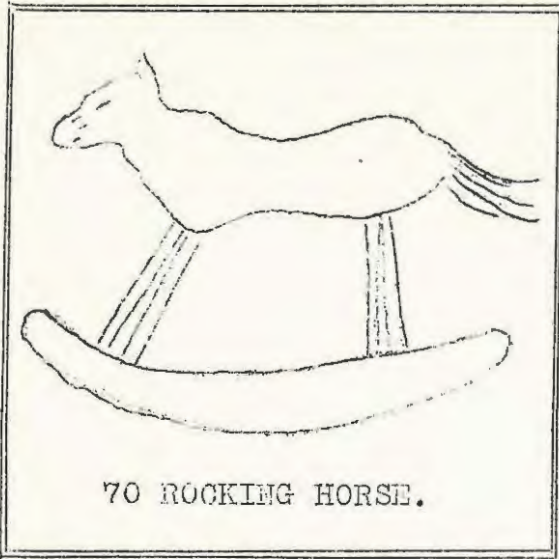
305 "Child's cart"
----, ----, B.



3550 "Cart"
Fo3, ---, BFo2.



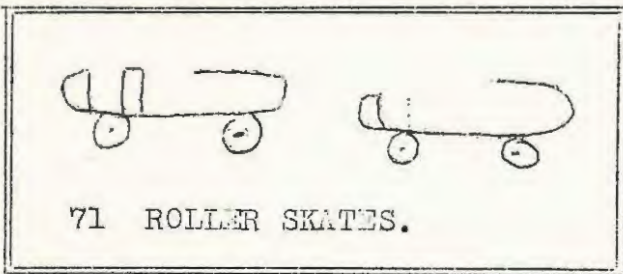
418 "Cart Wheel" Ho2, H2, Hox2.



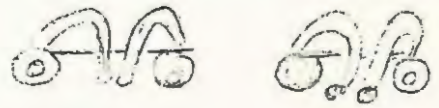
70 ROCKING HORSE.



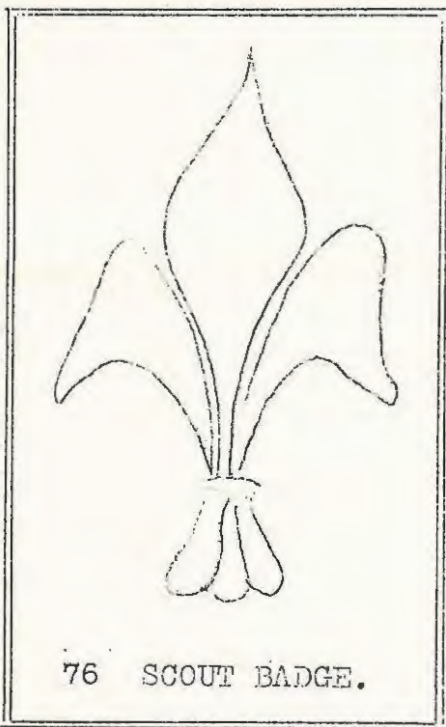
14237. "Rocking Horse"
AFo3, AF3, AF3.



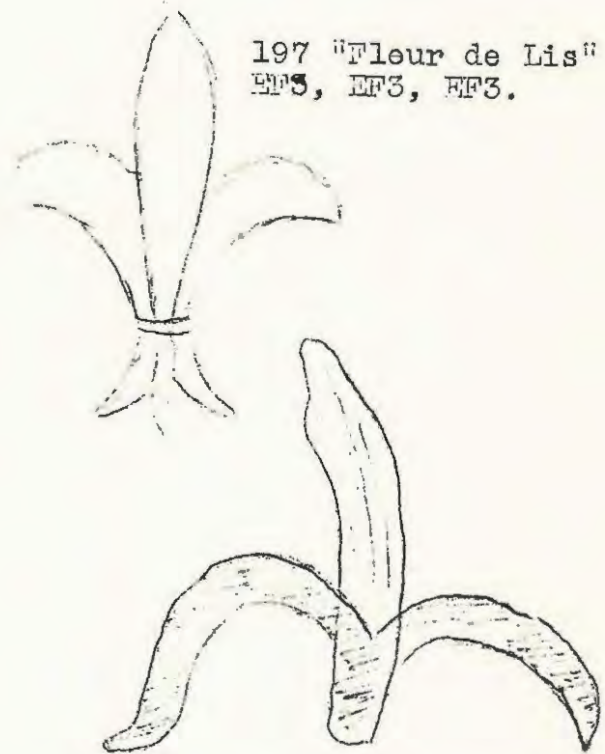
71 ROLLER SKATES.



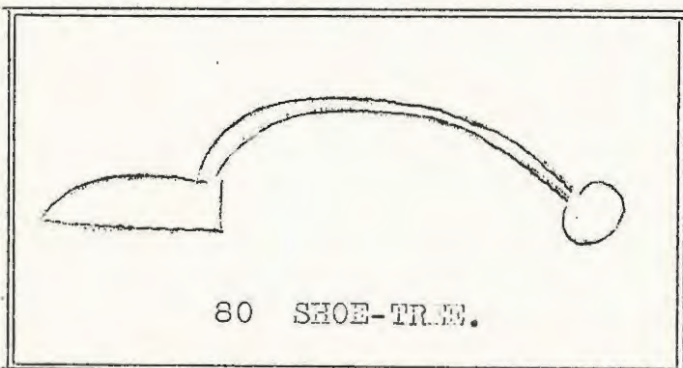
19178 "Skates"
BH2, BH2, BHx2.



76 SCOUT BADGE.



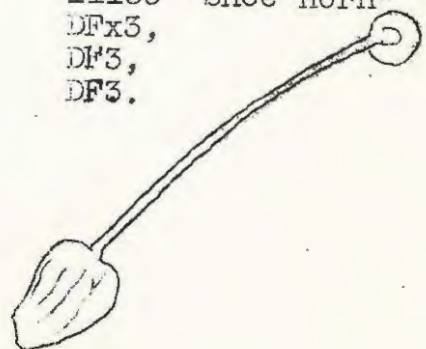
197 "Fleur de Lis"
EF3, EF3, EF3.



80 SHOE-TR.M.

1370 "Half-peeled banana"
Fx1, Fx2, Fx2

11135 "Shoe Horn"
DFx3,
DF3,
DF3.



(7) Classification of the subjects' drawings into
idea and shape resemblances:

In order to arrive at the data presented overleaf in tables XX and XXI, the writer systematically inspected each drawing appearing in the sample, and attempted to decide whether it satisfactorily represented the idea of the target, and whether it adequately reproduced the shape of the original. The bases for such judgements were admittedly subjective, but can be checked by the reader's own inspection of the data. A quick initial survey revealed the fact that there were a fair number of absolutely correct title or idea hits (the A hits), but that there were very few, if any, corresponding photographic likenesses in shape. Consequently a more lenient standard has been adopted for assessing shape than for idea similarities. Shape has been assessed either in terms of gestalt appearance, or by the presence of all the components shown in the target drawing.

Since more than one hit was scored on some originals it is possible that the same original can appear under different heads in the tables. In each case, however, the writer has first tried to make a general decision on all the hits scored on a particular original, before assigning it to more than one category. In some cases it was difficult to assign an original to one or the other category, perhaps because it showed neither good shape nor good idea resemblances. Such originals were assigned to category "d".

Table XX presents the data for the E.S.O., and the conclusions have been drawn from this data. The Control data is also given in table XXI, in case any reader might wish to refer to it.

These tables follow overleaf:

TABLE XX

CLASSIFICATION OF HITS ON E.S.O.
INTO IDEA AND SHAPE RESEMBLANCES.(a) Idea correct, but shape different:

- Anvil: Target squat, with small projecting tongue.
Reproductions have long tongues, conventional shape. (335)
- Hourglass: Target elongated and cylindrical, reproductions
either bulbous or semi-bulbous figure 8 shape. (334)
- Fire-extinguisher: Target cylindrical, reproductions all
conical (336).
- Bus stop: Target with round notice, reproduction with square. (337)
- Ace of Hearts: Target omits lines for edge of card, and small
hearts under A's. Reproductions supply both
these items. (338).
- Saint: Postures of reproductions entirely different. (339).
- Flat iron: Target non-electric. Reproduction with connector
and cord. (340).
- Dumb-bell: Target shows thin weights, cigar-shaped in end
section; reproductions show spherical weights. (341).
- Beacon: Target is light beacon, reproduction survey beacon. (344).
- Penguin: Shapes and postures of reproductions differ from
that of target. (345).
- T-square: Target horizontal and narrowing towards end.
Reproduction vertical and four-square. (345).
- Trident: Target shows no points to prongs, and gives them
rounded bottoms. Reproductions do not reproduce both
these items. (349).
- Tin opener: Very different shapes from that of target. (351).
- Rosette: Target shows three swallow-tail ribbons, the
reproductions two pointed ribbons. (353).
- Knobkerrie: Target spoon-shaped, reproductions with spherical
ends. (353).
- Violin bow: Position and detailed shape of bow are not shown. (357)
- Petrol Pump: Target shows no hose, all the reproductions do. (358).
- Saturn: Perhaps the most interesting examples of all. No
reproductions succeed in producing the original
completely, but 15716, 6196 and 15645 succeed in
reproducing the double ring in various distorted
forms. (360).
- Weathercock: Direction arms N.E.S.W. omitted from all
reproductions. (361).

Total Originals: 19.

(b) Cases where the shape has been reproduced, but the idea
mistaken:

- Vice: 14179 Spanner 14702 Saw. (337).
- Saint: 92 Rugby game; 3503 Telegraph poles; 2112 Braaivleis. (339)
- Penguin 10966 Pliers. (345)
- See-Saw: 3570 Open book; 11040 Balance; (347).
- Trident: 6333 Aloe; 12333 & 12135 Toasting fork; 7645 Gardening
fork; 10031 Fork; 19344 Pronged fork. (349)
- Bishop's Mitre: 4481 Church window; 7537 Church - lent. (355).
- Artist's Palette: 18131 Pattern. (356).
- Boomerang: 18606 Figure Seven. (359).

Total Originals: 8.

N.B. The reproductions referred to will be found at the
page reference given in brackets.

TABLE XX
(continued, page 2)

(c) Idea and shape both good:

- Snail: Some fairly close correspondences in shape, titles correct. (338).
Flit gun: Close correspondence in shape, except for conical nose. Title shows flit gun when true title is flit spray. (340).
Wish-bone: Close correspondences in shape, but not in position drawn. Titles differ only in word division. (341).
Spiffy: Good resemblance in title and shape. (342).
Volcano: Titles all correct. 14698 and 4713 very close correspondences in shape. (346).
See-Saw: Titles good: slight minor differences in shape and position of arm. (347).
Rolling Pin: Two good titles, 9721 and 3338 good shapes. (348).
Fez: Fairly close correspondences in title and shape, but very many elaborations in both. (350).
Whale: 8692 and 4048 very good shapes, remainder poor. Titles good. (354).
(Artist's) Palette: 14665 very good shape. No title includes (Artist's), otherwise good. (356).
Bottle Opener: Fair resemblances in shape, title good. (336).
Boomerang: Close resemblances in shape and title. (359).

Total Originals: 12.

(d) Cases difficult to classify: or idea & shape both bad:

- Vice (337).
Hoe (340).
Candlestick (342).
Mr. Chad (353).
Bishop's Mitre: (355).
Barbed Wire Barb (352).

Total Originals: 6

N.B. Some originals appear under more than one heading.
 The reproductions referred to will be found at the page references given in brackets.

TABLE XXI

CLASSIFICATION OF HITS ON C.S.O.
INTO IDEA AND SHAPE RESEMBLANCES.

(a) Idea good, shape poor:

Accordeon: Different perspective shown. (363).
Arc de Triomphe: Idea translated into English correctly, but shape poor: rounded instead of square top. (363).
Billiard Table: Idea correct, but different perspective shown.
Bellows: Titles correct, shapes poor. (364). (364).
Decanter: Totally wrong shapes, titles correct. (366).
Dust-pan: Correct title, but perspective different. (366).
Earphones: Ditto. (367).
Eiffel Tower: Ditto (367).
Father Christmas: Ditto (367).
Laurel Crown: Especially 14418: thistle & laurel wreath. (370).
Leopard: Idea fair, totally wrong shapes. (371).
Prickly Pear: Compare components of 17561 with target: all there but gestalt very different. (372).
Roller skates: Idea correct, but shape distorted. (374).
Shoe horn: Idea correct, but shape only fair. (374).

Total Originals: 14.

(b) Shape good, idea poor:

Ricksha: Note 3258 and especially 3550. (373)
Scout Badge: See 197 and 1370. (374).

Total Originals: 2.

(c) Shape and idea good:

Aladdin's Lamp: Spelling errors in title, but shapes rather good, particularly 13647, 3166 and 9823. (363).
Car jack: Shape very good when considered from point of view of components (base, angle of arm, etc.) (365).
Cupid: Shape good in curls, position of arms, bow etc, quiver omitted, idea correct. (365).
Discus Thrower (Model): Both shape & title close. (366).
Gollywog: Good shapes & titles. (368).
Hedgehog: ditto. (370).
Laurel Crown: 15413 good. (370).
Peacock: Especially 4318. (371).
Ricksha: Especially 1461, 3347. (373).
Rocking Horse: Title correct, shape good. (374).

Total Originals: 10

(d) Shape and idea poor, or difficult to classify:

Crutches: Singular instead of plural given: shape square tip instead of round, arm-rest concave instead of convex.
Gas mask: Different mask represented. (369). (365).
Harrow: Ditto. (369).
Hasp and Staple: Only depicted as minor details. (369).
Punch bag: The punch ball depicted is a different concept. (372)

Total Originals: 5.

N.B. Some originals appear under more than one heading.
 The reproductions referred to will be found at the page references given in brackets.

(8) Conclusions drawn from an examination of the sample of drawings presented, and from tables XX and XXI:

The material would appear to support the following generalizations:

(a) Very few, if any, of the good shape resemblances are photographic likenesses of the target originals.

(b) The best shape resemblances occur when the target is a very simple object of conventional design. If the agent sticks to this conventional shape when making his target drawing, fairly good reproductions result. (E.g. Spiffy, Palette (14665), Boomerang, Wish-bone and Flit gun etc.).

(c) When the agent deviates from the conventionally accepted pattern of the object, very few of the subjects reproduce the deviation in their drawings. (C.f. Hour glass, Fire-extinguisher, Petrol pump). There are occasionally fumbling attempts to do so which are very instructive. The hits on the target "Saturn" are of exceptional interest in this respect, and so will be discussed in detail.

The first point of interest is that Saturn was used as a target only in week E, on the second-last day of the experiment. The serial numbers of the hits scored on it thus show them all to be precognitive. The agent reports that when this title appeared out of the selector box, he found he had no idea of the appearance of Saturn and referred to a dictionary. He learnt that it was a planet with two rings and nine moons. When he drew it, he drew the rings co-axially, but not in the same plane, and inserted all nine moons, producing a highly distinctive version of Saturn, differing markedly from the conventional idea of this planet. The nine moons seem to have been too much for all the subjects, for they have been omitted from all the reproductions. Four of the subjects produced conventional drawings of Saturn, (11512, 10420, 4443, and 9658), with one ring, entitled "Saturn", "Saturn", "Moon" and "World". The other three (15716, 6196, and 15645) appear to

be groping towards the concept to be drawn later by the agent. Suggestions of a second ring are visible in 6196. In 15645 the subject has sketched a first ring with confidence, and then hesitantly dotted in a second ring above it, in the correct position shown by the target. Significantly she has refrained from calling her production "Saturn", but has contented herself with "Planet". The whole drawing suggests a hesitancy to accept the evidence of the E.S.P. impression she was getting, possibly because it went counter to her knowledge of the appearance of Saturn. The last example, 15716, is even more interesting. It is as if the subject had eked out his drawing from fragmentary and imperfect verbal communications from the agent. It is as if the agent had said: "I am drawing Saturn it has two rings, in different positions round it....." and at this point the communication channel broke down, and the subject was left to make the best drawing he could from the information which he had gleaned. He struggled with the position of the rings, but because his reproduction did violence to his concept of Saturn he titled it doubtfully "Planet ?", and not "Saturn" which he felt it could not be.

(d) It would seem justifiable to say that the above evidence refutes the idea that the subjects were directly apprehending the drawings in a clairvoyant way, or were obtaining their impressions through anything akin to a mental image. It would appear rather that E.S.P. in this experiment depended on the transmission of concepts (akin to verbal or ideational concepts), usually of a fragmentary nature. This confirms some of the findings of Warcollier (95).

(e) The drawings in the sample show many instances of elaboration and distortion of the impression in coming to consciousness, confirming Hettinger's findings (25). It is suggested that some of the forces tending to produce

these distortions are

- i) a tendency to conventionalize impressions;
- ii) a tendency to suppress;
- iii) a tendency to elaborate; and
- iv) a tendency to rationalize.

(f) The results also demonstrate the fitful and unreliable nature of the E.S.P. process. A fair proportion of the E.S.C. had no A hits scored on them whatsoever. This raises another interesting point for further investigation: what constitutes a "favourable" target original, and what renders a particular target "difficult". The reader might like to ponder the following list of E.S.O. on which no A hits were scored, and try and find some common element. Such a constant has not been discovered by the writer. The titles are: Morse Key; Flit spray; Turnstile; Javelin; Pendulum; Asterisk; Barbed Wire Barb; U-tube; Sentry box; Wing nut; Honeycomb; Cricket pads; Taj Mahal and Rhinoceros.

(g) It will be remembered that the Linkage findings suggested that some interaction was taking place between agent and subjects. This qualitative analysis has shown that the interaction is in the nature of a transmission of concepts, and hence would appear to be telepathic rather than clairvoyant in substance, at least under conditions of this experiment. This confirms and extends the statistical findings previously obtained.

This brings us to an end of the data which will be presented in this thesis, and it remains next to integrate and summarise the findings from all the analyses that have been carried out.

GENERAL CONCLUSIONS DRAWN FROM AN INTEGRATION
OF ALL THE FINDINGS.

In the course of the account that follows it is intended to deal first with the evidence against any particular hypothesis, and then with the evidence in its favour. It will be remembered that the first stated aim of the investigation was to discover whether any facts supported the view that an extra-sensory exchange had taken place between subjects and agent which made it possible for them to gain paranormal knowledge of the target drawings made by him. For brevity let us refer to this as the Telepathy Hypothesis.

At first sight there are two findings from the experiment that weaken the Telepathy Hypothesis. The first of these is that in the hit categories allowing fainter and more distorted resemblances in title, shape and association, the subjects have sometimes returned more hits on the C.S.O. than on the E.S.O. It was pointed out in explanation of this that the two groups of originals had only been equated in terms of the strict A type hits. It was suggested that the observed differences here were due to the initial inequality of the two groups of originals in terms of these fainter types of hit, and hence not disproof of the telepathy hypothesis. It was also pointed out that these categories of hit were the ones most open to chance dilutions.

The second effect was that the hits scored on any original tended to distribute evenly throughout all five weeks of the experiment, and showed no tendency for the most hits to be scored on an original in the week when it actually was a target than in other weeks. If chance alone were producing the hits, this is the pattern one would expect to occur. Moreover, Whately Carington had previously reported that his results

showed a tendency for most hits to be scored on his originals at the time they were actually being used as targets, than in other experiments (13). Whately Carington had, however, used different subjects and different targets for each of his experiments.* In the Rhodes Experiment the same subjects participated in all five sections. It was suggested that the two procedures were not identical, and that this difference could account for the fact that the Rhodes results did not follow the pattern reported by Carington.

No outstandingly successful subjects were found among those tested, the number of hits scored falling into a normal probability curve. Hits on particular originals also tended to distribute normally, except that in the E.S.O. 14 titles were found distributed randomly throughout the sections, on which no A hits were scored whatsoever. These two facts are neutral, they neither disprove nor help to establish the Telepathy Hypothesis.

The investigation disclosed no other facts contrary to the hypothesis that some paranormal action was taking place between agent and subjects which enabled the subjects to score hits on the targets.

The fact that improving the linkage between agent and subjects significantly increased the number of hits scored by them renders untenable any other hypotheses, such as that the results were due to sensory leakage, or bias on the part of the judges, or were produced by any other normal or paranormal action of the agent, or of the experimenter, or that the subjects were gaining their knowledge normally or paranormally from the experimenter, or by precognitive telepathy or clairvoyance from this report itself. The linkage findings clearly establish some relation between subjects and agent

* In one instance, when he ran two experiments concurrently, he found that his subjects scored hits on the targets of both experiments indiscriminately. See p. 159 above. This is akin to the effect obtained in the Rhodes Experiment.

as responsible for at least part of the hits scored.

Other evidence in favour of the Telepathy Hypothesis is supplied by findings from the various analyses. Thus in the strictest hit category, the A hits, the subjects showed a highly significant positive deviation from empirically established chance expectancy on the E.S.O. actually used as targets in the experiment, but no significant deviation from the theoretically expected chance expectancy in the hits scored on the C.S.O. Examination of the personality material disclosed that the high-scoring subjects were significantly more extraverted than the low-scoring group, a relationship previously discovered by independent investigators. It was also found that the data showed other relationships between personality traits and hit scoring ability that had been predicted on the basis of research findings previously reported. Finally, when correct linkage was supplied to the Experimental Group Subjects, designed to link them more closely to their agent, and thus help them score hits, they showed a significant increase in scoring rate. A Control Group of Subjects, receiving equated but incorrect linkage, showed no such consistent significant fluctuations. It would be very hard to explain the consistency of these results as due to the operation of uncontrolled chance factors even with special pleading.

There is another type of evidence which also tends to confirm that the results are genuinely due to the subjects exercising E.S.P. ability. It is of this nature. Generally speaking the strict hit categories produce the clearest evidence in support of the E.S.P. hypothesis. As less strict hit categories are used, more open to dilutions from chance factors, the results become progressively less clearcut. A similar indication of internal consistency is supplied by the Bernreuter Introversion scores. It is found that if the subject groups are arranged in hierachical order in terms of the E.S.P.

A hits scored, their mean introversion scores fall into the same hierachical order. If the results observed were merely due to chance, there is no reason why they should be expected to show this logically consistent pattern. We conclude, therefore, that the results of the experiment demonstrate the operation of an E.S.P. process between agent and subjects.

We can turn next to the second stated aim of the investigation, namely, to find whether any relationships exist between the subjects' hit scoring ability and their personality makeup. Five predictions were made as to what trends the data should show if it were genuinely being produced by the subjects' E.S.P. powers. As only the relationship with extraversion could be demonstrated with satisfactory statistical significance it might be argued that this section supplied but meagre support of the predictions. It should be pointed out, however, that in three out of the four remaining trends the deviations were in the predicted direction, and that previous investigators had pointed out that the effects could be expected to be very small. (65, page 25). Had considerably more subjects taken part in the Rhodes experiment, it is possible that these trends might have reached statistical significance. Moreover, rather different measures of the various factors such as psychological adjustment were used in the Rhodes Experiment than were used by the originally reporting research workers. We conclude that the personality findings contribute additional (though not too strong) support of the paranormality of the Rhodes data. The results from the Psychic Questionnaire should be regarded as only suggestive, since they fail to reach statistical significance.

Finally, let us consider whether the experiment gives a precise answer to the last problem enunciated in the aims, namely, whether supplying the subjects with linkage material designed to connect them with their agent increased

their ability to score hits.

The first assessment carried out on pooled data from weeks A, B, C, D & E, did not permit any conclusions to be drawn, for it showed significant fluctuations in scoring rate for both subject groups on both series of originals. The hypothesis was suggested that these fluctuations were caused by the hindrance effect of the colour square linkage in week B. When this was tested by excluding week B scores, and assessing the pooled data of weeks A, C, D & E only, the hypothesis was confirmed, for significant fluctuations were found only in the E.G.S' hits on the E.S.O., a result consistent with the validity of the linkage hypothesis.

When a third type of assessment was made, in which the weekly scoring rates under different conditions of linkage were separately compared with week A in which no linkage was supplied, 16 fluctuations reaching the .02 level of significance (two tail) were found in a table of 104 assessments. Eleven out of sixteen of these occurred in respect to improvements shown by E.G.S. hits on E.S.O., three were diminutions in respect of E.G.S. hits on C.S.O. in week B, and the remaining two significant changes occurred in respect to C.G.S. hits on E.S.O., one diminution in week B, and one increase in week E. These last two significant changes have been ignored in assessing the results on the grounds that they are atypical, and in a table of 104 calculations could reasonably be regarded as chance produced variations. The attention of the reader is directed to them merely so that he can weigh the ethics of disregarding them when arriving at the conclusions. The negative deviations in week B in the E.G.S. hits on the C.S.O. can be attributed to the hindrance effect exerted by the colour square links, in addition to any correctness factor effect. The finding that eleven out of the sixteen significant fluctuations occur in the E.G.S. hits on E.S.O., and are all positive, strongly supports the hypothesis that some correct linkage increased subjects' hit-scoring ability.

The fact that the Scores Corrected for Bunching give less clear-cut results might be taken by some critics as weakening the evidence for the Linkage Hypothesis. In defence of this hypothesis it might be said that they do not contradict the findings from the Raw Score assessments, but tend to confirm them. More weight was attached to the Raw Scores because even if they were in part produced by uncontrolled factors, there is no reason why chance alone should cause such factors to operate selectively only on the E.G.S. hits on the E.S.O. It was felt that the Scores Corrected for Bunching were based on a crude empirical procedure based on inspection rather than on any exact statistical procedure, and that these corrections were applied to test a specific hypothesis in quite another connection than the present one.

All the possible arguments against the Linkage Hypothesis seem to the writer to be trivial. The case for the hypothesis would appear to be very strong. The only significant and consistent improvements in score took place in hits on the E.S.O. when the E.G.S. were supplied with handkerchiefs actually used by the agent, photographs of him, and a description of his personality. No other such significant increases were observed in the case of the E.G.S. hits on the C.S.O. It would seem safe to conclude that uncontrolled factors were unable to account for the significant improvements in scoring rate shown by the E.G.S. on the E.S.O. Once the validity of these findings is admitted, further conclusions can be drawn about the way in which the linkage material acted in producing its effect.

Firstly, the results point clearly to some sort of interaction taking place between the minds of the subjects and that of the agent - that is, a telepathic process in the main, and not predominantly a clairvoyant one. This is supported also by the qualitative survey undertaken in the

Fifth Main Analysis. It will be remembered that this analysis suggested that the subjects results were achieved by apprehending ideas and concepts, and not by any sort of direct clairvoyant apprehension of the targets themselves. These results tended to bear out the previous findings of Warcollier (95), and especially Osty in (93) (See page 276 above).

Secondly, the linkage material appears to have worked by putting the subjects in contact with the agent's mind as a whole, and not to specific items in it. This evidence seems to disprove Whately Carington's Association Theory of Telepathy (14), which postulates such an item to item relationship between the minds of agent and subjects.

Thirdly, it was found that the hits on any particular original tended to be distributed evenly throughout the whole period of the experiment, another argument against the Association Theory.

Fourthly, it was found that hit-scoring ability was correlated with the subjects degree of extraversion as measured by the Bernreuter Personality Inventory Scale.

Fifthly, it was demonstrated that the effectiveness of the linkage material depended on its correctness, at least in part. The experiment contributed no information as to the possible existence and contribution of a suggestion factor to the effectiveness of the linkage.

Finally, the material from the Fifth Main Analysis tended to support the view that the telepathic impression was subject to distortions, omissions, elaborations and fragmentations in coming to consciousness, and that the operation of E.S.P. ability tended to be inconsistent and favour some experimental target originals more than others. These findings are in line with those reported by Hettinger (25).

Summing up in terms of the originally stated aims, we can conclude that the subjects demonstrated the existence of E.S.P. ability, that this ability was correlated with certain personality traits, and that supplying the subjects with certain types of linkage increased their scoring rates.

FINAL REFLECTIONS AND
SUGGESTIONS TO OTHER EXPERIMENTERS
WANTING TO REPEAT THIS EXPERIMENT.

During the course of the assessments made in this investigation each new fact revealed has been carefully considered, and an explanation sought for it in terms of normal and paranormal causes. As more and more facts accumulated from the various analyses, a gestalt pattern began to emerge which was logical and consistent with the E.S.P. Hypothesis, but not with any of the alternative hypotheses which came to mind. This pleasing logical consistency of the results is, to the writer's way of thinking, a powerful additional argument for the acceptance of the E.S.P. and Linkage Hypotheses, over and above the mere statements of statistical probability. Part of the difficulty confronting a parapsychologist is that the occurrence of E.S.P. can only be established by eliminating all other counter hypotheses, and there always remains the possibility, however slight, that a normal explanation capable of accounting for the results has been overlooked. To deny the existence of such a possibility is to commit the sin of hubris. Nevertheless, in view of the lawful effects which the data shows, and its internal consistency, the writer would personally be surprised to find that any factors capable of accounting for the results obtained in normal terms had been missed. One of the aesthetically pleasing rewards from undertaking this research has been to see this meaningful pattern emerging from the results of the constituent analyses. This is not to say that the data has always fulfilled a priori expectations, but it has shown a number of the trends previously established as being typical of E.S.P. data. The results of the Rhodes Experiment thus offer strong evidence for the existence of E.S.P. and the effectiveness of certain linkage procedures. They also confirm many of the reported findings of previous investigators. With respect to the new facts which the

Rhodes Experiment appears to have established regarding linkage, as with all research findings, it would be of great value for them to be checked by independent investigators in other experiments. The writer personally feels that it would be better to test the findings herein reported by different procedures specifically designed to overcome the weaknesses and inherent limitations of the Rhodes Experiment. Some workers may desire to repeat the procedure of this experiment as it stands, in which case the following advice may be useful.

- (1) Run a small pilot experiment to discover possible snags before embarking on the main experiment.
- (2) Use considerably larger numbers of subjects.
- (3) Select the target originals carefully to avoid confusions of shape and title.
- (4) Pay attention to the morale of the subjects, and keep up their interest in the experiment by personal contacts with them throughout the experiment.
- (5) Check for omissions and errors step by step, particularly in any work done by helpers.
- (6) If subject numbers are sufficiently big, have a third group of subjects who are given no linkage throughout the experiment as an additional Control Group.
- (7) Record all data on Hollerith cards, and arrange for machine scoring of the personality test material.
- (8) Obtain statistical advice from recognised statistical authorities as to the types of assessment to use.

Many of these suggestions sound simple and obvious, but failure to formulate them early in the investigation cost the writer much unnecessary labour. Expense rendered others beyond the

reach of the Rhodes investigation.

For those wishing to break new ground, and extend the Rhodes findings, the following problems are suggested as being worthy of further investigation:

- (1) Does linkage exert a cumulative effect ?
- (2) Does a suggestion factor contribute to the effectiveness of the linkage material ?
- (3) What is the relative importance of the correctness factor compared with other operative factors contributing to the effectiveness of the linkage material ?
- (4) Why are no hits scored on certain of the target Originals ?

Many more problems could be cited, but the above will give some indication of the very interesting possibilities for further investigation in the parapsychological field.

GLOSSARY OF TERMS AND ABBREVIATIONS

USED IN THE REPORT.

In this report technical terms have been used in the sense defined in the Glossary given in "The Journal of Parapsychology", whenever possible. Many of the definitions given below come directly from this source.

Agent: The individual who attempts to transmit the target in a telepathy experiment.

A Hit: Classification awarded when the title of a subject's drawing is identical with the title of the agent's target. E.g. "Fountain" for "Fountain".

Association Theory of Telepathy, Whately Carington's: The theory that if an agent associates a target with a particular idea, and this idea is present also in the mind of a subject, the association is automatically formed in the mind of the subject as well, and results in the subject becoming aware of the target. Such common ideas are called K ideas. (13)
See pages 22 & 278 above.

B Hit: Classification awarded when a subject generalizes the title of a target, e.g. "Mug" for the target "Beer Mug".

Bunching Errors: A-typical spurts of hits on a particular original occurring at a time when some external event suggests the target to the subjects. See page 116.

C.R. = Critical Ratio: A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The C.R. is obtained by dividing the observed deviation by the standard deviation. (The probability of a given C.R. may be found by consulting tables of the probability integral, such as Pearson's)

C.G.S. = Control Group Subjects: The group of subjects who received incorrect linkage material in the Rhodes Experiment.

C Hit: Classification awarded when a subject particularises a target title, e.g. "Beer Mug" when the target title is "Mug".

Clairvoyance: Extrasensory perception of objective events as distinguished from telepathic cognition of the mental activities of another person.

Communicator: An entity communicating through, and claiming to be independent of a medium.

Compressive type: A subject who, when making drawings, uses the space at his disposal timidly and hesitantly. (Elkisch, 19).
See page 173 above.

Control: (1) The personality which takes charge of the medium's body when the normal personality is in abeyance.
(2) Experimental procedure designed to provide a check on the influence of uncontrolled factors in an investigation.

C.S.O.= Control Series Originals: The 50 drawings not used as targets in the Rhodes Experiment, against which the subjects' drawings were assessed to determine the numbers of hits which could be expected by chance with no E.S.P. operating.

D as a statistical symbol: Deviation. The amount an observed number of hits or an average score deviates from the mean chance expectation or chance average.

D hit: Classification awarded when a subject gets the general name of a target original right, but not the specific species. E.g.: "Beer Mug" when the target title is "Tooth Mug".

Decline effect: The tendency for the number of hits scored in an E.S.P. experiment to be greatest at the beginning of the experiment, and then to fall off progressively until the end is almost reached, when they sometimes show a terminal spurt. This is also known as the J-curve effect.

Dirt: Term used by the Arapesh for emanations from a person's body, or relics of it, thought by them to be effective linkage material for purposes of sorcery, by virtue of the laws of contagious magic. See page 271 above.

Displacement: E.S.P. responses to targets other than those for which the trials were intended. Backward displacement is responses to targets preceding the intended targets, forward displacement is responses to targets coming later than the intended targets.

E.G.S. = Experimental Group Subjects: Subjects receiving correct linkage material in the Rhodes Experiment.

E Hit: Classification awarded when subject gives a synonym for actual target title, e.g. "Car" for "Automobile".

Empirically determined Chance Frequency: The mean Raw A hits scored in the Main Experiment on a sample of 200 series of 50 originals selected at random from all the titles suggested only once in the Preliminary Frequency Experiment. See page 146 above.

E.S.O. = Experimental Series Originals: The 50 drawings made as targets by the agent in the Rhodes Experiment.

E.S.P. = Extrasensory Perception: awareness of, or response to an external event or influence not apprehended by sensory means.

E.S.P. Cards: Cards, each bearing one of the following five symbols: star, circle, square, cross, and waves. A pack consists of 25 cards, five of each suit. The cards are used as target material in E.S.P. research.

Ethereal (Light): Vague concept offered by Mrs. Piper's controls as a title for the effective force of an influence. See page 274 above.

Expansive type: A subject who, when making drawings, uses the space at his disposal boldly. (Elkisch, 19) See page 173 above.

Experimenter: The person who administers an experiment to subjects, and who usually devises and organises it, and interprets the results. In the Rhodes Experiment the writer of the report was the experimenter.

Exposure Box: A light-tight locked box, in which the agent put pairs of originals during the 24 hours that they were the experimental targets in the Rhodes Experiment.

Extrapunitive Type: An individual who turns his aggression outward onto the environment. (Rosenzweig) (68) See page 183 above.

Extravert: One whose attention is given mainly to the external environment rather than to his own mental processes.

F Hit: Classification awarded when subjects reproduced the shape of the target drawing without omissions or additions.

Frazer's Law of Contagious Magic: "Things that have once been in contact with one another, continue to act on each other at a distance after the physical contact has been severed" (21) See page 272 above.

Fully Corrected A hits: The residue of A hits remaining after all doubtfully valid A hits have been eliminated. See page 128 above.

G.E.S.P. = General Extrasensory Perception: Extrasensory Perception permitting either telepathy or clairvoyance or both to operate.

G. Hit: Classification awarded when only part of an object was drawn as a target by the agent, but the whole object was reproduced by the subject.

Goat: Term applied by Schmeidler to a subject sceptical of the existence of E.S.P. (See 60 to 70 in Bibliography) See page 179 above.

Greville's Hypothesis: The hypothesis that the true chance frequencies for scoring A hits are less than those suggested by the Preliminary Frequency Experiment. See page 144 above.

H Hit: Classification awarded when subjects reproduce only part of a target drawing.

I Hit: Classification awarded when there is a clear association between a subject's reproduction and the target drawing, but the name and shape are different. E.g. "Wind" for "Weathervane".

Impunitive Type: An individual who suppresses & glosses over his aggression. (Rosenzweig, reported by Schmeidler, 68). See Page 183 above.

Influences: Articles held by a medium as an aid to the mediumistic process, or as a guide to making contact with a desired communicator. See page 273 above.

Interval Scale: A scale in which the distance between any two numbers on the scale are known, in addition to the scale having all the characteristics of an ordinal scale. (See Siegel, 107, page 25-6)

Introvert: One whose attention is given mainly to his own mental experiences rather than to the external environment.

J Hit: Classification awarded when a subject reproduced the right idea, but the name and shape are wrong. E.g. "Chrysalis of Man" for the target "Mummy"

J Curve: See decline effect.

Judge: Official who compared the subjects drawings with the E.S.O. and C.S.O. targets, assessed the resemblances shown, and awarded hits according to the classification scheme provided.

K Hit: Classification awarded for miscellaneous resemblances not covered by any of the other hit categories.

K ideas: Ideas present in the minds of both agent and subject in an E.S.P. experiment which are associated with the targets, and which mediate the telepathic interaction. (Whately Carington) (13) See also page 278 above.

Law of Frequency: The law which states that other things being equal, the more frequently a particular association occurs the more likely it is to come to mind subsequently.

Law of Recency: The law which states that other things being equal, recently formed associations will tend to come to mind more readily than associations formed earlier.

Law of Vividness: The law which states that other things being equal, the more vividly an association is formed, the easier it will be to recall afterwards.

Light (Ethereal): See ethereal.

Linkage Hypothesis: The hypothesis that material once associated with the agent in the Rhodes Experiment, and supplied to the subjects, effectively increases their hit-scoring ability.

Linkage material: Material in some way connected with the agent in the Rhodes Experiment, and supplied to the subjects in the belief that it will create bonds between them and the agent, and facilitate the E.S.P. process.

Linkage problem: The complex of unknown factors and processes involved in an E.S.P. action between subjects and agent, which results in the subjects selecting, from all the possible agents and targets available, the appropriate elements to their particular needs. (See page 270 above).

Main Experiment: The part of the Rhodes Experiment carried out after the Preliminary Frequency Experiment, and designed to investigate the effectiveness of linkage material.

M as a statistical symbol: The mean.

Medium: Person claiming to be an instrument used by non-natural entities to communicate with and influence the mundane world.

Metagnomy: Synonym coined by Osty for psychometry, to avoid the confusion due to the psychological meaning of the latter term.

N as a statistical symbol: Number of cases.

Nominal Scale: "Measurement at its weakest level exists when numbers or other symbols are used simply to classify an object, person or characteristic. When numbers or other symbols are used to identify the groups to which various objects belong, these numbers or symbols constitute a nominal or classificatory scale. (Siegel: 107, page 22).

O (1) As a symbol in the scoring code: Awarded when the subject has introduced elaborations on the target drawing.

(2) As a symbol in statistics: Observed frequency.

Ordinal Scale: or ranking scale: "Given a nominal scale of equivalence classes, if the relationship "greater than" holds for all pairs of classes, the scale is an ordinal scale." (Siegel, 107, page 23-24).

Original: A drawing used either as an experimental target, or as a control target.

Paranormal Cognition: Synonym for Extrasensory Perception.

Parapsychology: A division of psychology dealing with behavioral or personal effects that are demonstrably nonphysical (that is, which do not fall within the scope of known physical principles).

P as a statistical symbol: Probability: A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

Psi missing: Exercise of Psi ability in a way that avoids the target the subject is attempting to hit.

Percipient: The subject in an E.S.P. experiment who attempts to gain a Psi impression from the agent or target.

P.F.E. = Preliminary Frequency Experiment: The experimental procedure used in the Rhodes Experiment to establish the A hit chance frequencies of possible target originals, and enable the E.S.O. and the C.S.O. to be equated in difficulty.

P.K. = Psychokinesis: The direct influence exerted on a physical system without any known intermediate physical energy or instrumentation.

Pre cognition: Cognition of a future event which could not be known through rational inference.

Post cognition: Cognition of a past event coming through channels other than the normally accepted senses.

Psi. A general term to identify personal factors or processes which are non-physical in nature. It approximates the popular use of the word "psychic" and the technical one "parapsychical"

Psychical Research: Older term used for parapsychology.

Psychometry: Term coined by Buchanan (11), to refer to the process whereby sensitives supposedly obtain their knowledge supernormally from an object held in their hands.

Previously Determined Chance Frequency: The frequencies of scoring A hits established by the Preliminary Frequency Experiment.

Raw Scores: Uncorrected scores derived from the assessments of the three primary judges in the Rhodes Experiment, without additions or deletions.

Retro cognition: Synonym for post cognition.

Rhodes Experiment: The experimental investigation forming the subject matter of the dissertation "Linkage in Extrasensory Perception", carried out in 1947.

S.D. = Standard Deviation: Usually the theoretical root mean square of the deviations.

Selector box: Box in which the 100 title slips were put and from which the 50 E.S.O. titles were shaken at random by the agent, leaving 50 titles in the box later to form the C.S.O.

Sensitive: Individual with more than average control over his or her E.S.P. ability.

Sheep: Technical term coined by Schmeidler (60 to 70), and designating a subject in an E.S.P. experiment who accepts the reality of E.S.P.

Singles Title: A title listed by the analysers of the Preliminary Frequency Experiment data as having occurred only once in the 15,835 suggestions returned by the Rhodes subjects.

Spontaneous cases: Psi phenomena occurring spontaneously in nature, as contrasted with experimentally induced phenomena.

Statistical symbols: σ or S.D. = The Standard Deviation.

Σ

= The sum of.

$|x|$

= The absolute value of x without regard to sign.

$>$

= Is greater than.

$<$

= Is less than.

Subject: (1) The percipient in an E.S.P. experiment. (Normal usage).
(2) The person about whom a sensitive is attempting to obtain paranormal knowledge. (Hettinger).

Target: The drawing, idea, or object which the subject is attempting to apprehend paranormally.

Telepathy: Extrasensory perception of the mental activities of another person. (As distinguished from clairvoyant cognition of objective events).

Telepathy Hypothesis: The hypothesis that the observed hits in the Rhodes Experiment were at least in part caused by the subjects' successful use of E.S.P.

Ultra-perceptive faculty: Hypothetical structure in an individual mediating the E.S.P. process. (Hettinger, 25).

Winners: Hits scored on an original during the period it was actually being used as a target, as distinct from hits scored on it in other periods. (Whately Carington, 13).

In the Rhodes Experiment, winners are hits scored on an original during the actual week during which it was used as a target.

χ^2 = Chi Square: A sum of quantities each of which is a deviation squared divided by an expected value. Also the sum of the squares of Critical Ratios (C.R.'s).

BIBLIOGRAPHY.

(Abbreviations have been taken from)
the World List.

- (1) Allison, L.W.: Note on the repetition of Whately Carington's Experiments. J. Amer. Soc. Psychic. Res. 1945, 39, pp 210-242.
- (2) Allport, G.W.; and Allport, F.H.: A-S Reaction Study. Houghton Mifflin Co., U.S.A. 1928.
- (3) Allport, G.W., and Vernon, P.E.: A Study of Values. Houghton Mifflin Co., Boston, 1931.
- (4) Bates, K.E., and Newton, M.: An Experimental Study of E.S.P. Capacity in Mental Patients. J. Parapsychol. 1951, 15, pp 271-277.
- (5) Bell, M.: The Normal Function and Manifestations of Psi. J. Amer. Soc. Psychic. Res. 1952, 56, pp 96-106.
- (6) Bendit, L.J.: Paranormal Cognition. Faber & Faber, London, 1944.
- (7) Bernreuter, R.G.: The Personality Inventory. Stanford University Press, California, U.S.A., 1935.
- (8) Bevan, J.M.: E.S.P. Tests in Light and Darkness. J. Parapsychol. 1947, 11, pp 76-89.
- (9) Bevan, J.M.: The relation of attitude to success in E.S.P. scoring. J. Parapsychol. 1947, 11, pp 296-309
- (10) Broad, C.D.: Normal Cognition, Clairvoyance and Telepathy. Proc. Soc. psych. Res. 1935, XLIII pp 397-438.
- (11) Buchanan, J.R.: Manual of Psychometry. F.J. Hodges, Boston, 1893.
- (12) Buros, O.K. (Ed.): The Nineteen Forty Mental Measurements Yearbook. Mental Measurements Yearbook, New Jersey, 1941.
- (13) Carington, W.: Experiments on the Paranormal Cognition of Drawings. Proc. Soc. psych. Res. 1940-1941. XLVI, pp 34-151.
- (14) Carington, W.: Telepathy. Methuen & Co. Ltd., London, 1945.
- (15) Casper, G.W.: A further study of the relation of Attitude to success in E.S.P. scoring. J. Parapsychol. 1951, 15, pp 139-145.
- (16) Devereux, G. (Ed.): Psychoanalysis and the Occult. International Universities Press, New York, 1953.
- (17) Ehrenwald, J.: Telepathy and Medical Psychology. George Allen & Unwin, Ltd., London, 1947.
- (18) Eilbert, L.: A study of the moods and attitudes that operate to facilitate or retard the ability to score on a standard E.S.P. test. Unpublished Master's Thesis, School of Education, City College of New York, 1948.

- (19) Elkisch, P.: Children's Drawings in a Projective Technique. Psychol. Monogr., 1945, 58, pp 1-31.
- (20) Fraser-Nicol, J., and Humphrey, B.M.: The Exploration of E.S.P. and Human Personality. J. Amer. Soc. Psychic. Res., 1953, XLVII, pp 133-178.
- (21) Frazer, J.: The Golden Bough (Abridged edition). MacMillan & Co., Ltd., London, 1950.
- (22) Freud, S.: New Introductory Lectures on Psychoanalysis. Hogarth Press, London, 1946.
- (23) Grela, J.J.: Effect on E.S.P. scoring of hypnotically induced attitudes. J. Parapsychol., 1945, 9, pp 194-202.
- (24) Heston, J.C.: Heston Manual and Personal Adjustment Inventory. World Book Co., Yonkers on Hudson, 1949.
- (25) Hettinger, J.: Exploring the Ultra-Perceptive Faculty. Rider and Co., London, 1941.
- (26) Hodgson, R.: A record of observations of certain phenomena of Trance. Proc. Soc. psych. Res. 1892, VIII, pp 21.
- (27) Humphrey, B.M.: An exploratory correlation study of personality measures and E.S.P. scores. J. Parapsychol. 1945, 9, pp 116-123.
- (28) Humphrey, B.M.: Success in E.S.P. as related to form of response drawings: I: Clairvoyance Experiments. J. Parapsychol., 1946, 10, pp 78-106.
- (29) Humphrey, B.M.: Success in E.S.P. as related to form of response drawings: II: G.E.S.P. Experiments. J. Parapsychol., 1946, 10, pp 181-196.
- (30) Humphrey, B.M.: Handbook of Tests in Parapsychology. Duke University Press, Durham, 1948.
- (31) Humphrey, B.M.: The relation of E.S.P. to mode of drawing. J. Parapsychol., 1949, 13, pp 31-47.
- (32) Humphrey, B.M.: Further work of Dr. Stuart on interest test ratings, and E.S.P. J. Parapsychol., 1949, 13, pp 151-165.
- (33) Humphrey, B.M.: E.S.P. Subjects rated by two measures of Personality. J. Parapsychol., 1949, 13, pp 274-291.
- (34) Humphrey, B.M.: A New Scale for Separating High and Low-scoring Subjects in E.S.P. Tests. J. Parapsychol., 1950, 14, pp 9-23.
- (35) Humphrey, B.M.: E.S.P. Score Level Predicted by a Combination of Measures of Personality. J. Parapsychol. 1950, 14, pp 193-206.
- (36) Humphrey, B.M.: The relation of some personality ratings to E.S.P. scores. J. Soc. psych. Res. Lond. 1951, 36, pp 454-466.
- (37) Humphrey, B.M.: Introversion-Extraversion ratings in relation to scores in E.S.P. tests. J. Parapsychol. 1951, 15, pp 252-262.

- (38) Hunter, M.: Reaction to Conquest. Oxford University Press, London, 1936.
- (39) Kahn, S.D.: Studies in Extrasensory Perception Experiments utilizing an Electronic Scoring Device. Proc. Amer. Soc. Psychic. Res., 1952, XXV, pp 1-48.
- (40) Kohler, M.: The IZANCOMA Diviners. Government Printer, Union of South Africa, Pretoria, 1941.
- (41) Krämer, H., and Sprenger, J.: Malleus Maleficarum Trans. Summers, M.: Pushkin Press, London, 1948.
- (42) MacFarlane, J.D.: Discrimination shown between experimenters by subjects. J. Parapsychol., 1938, 2, pp 160-170.
- (43) Mead, M.: Sex and Temperament in Three Primitive Societies. George Routledge & Sons, London, 1935.
- (44) Mundle, C.W.K.: Selectivity in E.S.P. experiments. J. Soc. psych. Res. Lond., 1951, XXXVI, pp 385-393.
- (45) Murphy, G.M.: Psychical Research and Personality. Proc. Soc. psych. Res. Lond., 1949, XLIX pp 1-15.
- (46) Murray, M.A.: The Witch-cult in Western Europe. Oxford University Press, Oxford, 1921.
- (47) Oram, A.T.: An experiment with random numbers. J. Soc. psych. Res. Lond., 1954, XXXVII, pp 369-377
- (48) Pederson-Krag, G.: The Medical Section: The first three years. J. Amer. Soc. Psychic. Res., 1951, XLV, pp 115-116.
- (49) Pratt, J.G., and Price, M.M.: The experimenter-subject relationship in tests for E.S.P. J. Parapsychol. 1938, 2, pp 84-94.
- (50) Pressey, S.L.: Pressey X-0 Tests. C.H. Stoelting Co., Chicago, 1920.
- (51) Price, H.H.: Survival and the idea of "Another World". Proc. Soc psych. Res., Lond., 1953, 50, pp 1-25.
- (52) Prince, W.F.: Presidential Address. Proc. Soc. psych. Res. Lond., 1930-31. XXXIX. p 277
- (53) Reeves, M.P.: Whately Carington's Theory of Paranormal Cognition and Allied Phenomena: A Review. J. Amer. Soc. Psychic. Res., 1945, 39, pp 95-112.
- (54) Rhine, J.B., and associates: Extra-sensory Perception after Sixty Years. Henry Holt & Co., New York, 1940.
- (55) Rhine, J.B.: Psi Phenomena and Psychiatry. Proc. R. Soc. Med., 1950, XLIII, pp 804-814.
- (56) Rhine J.B.: The Problem of Psi-Missing. J. Parapsychol. 1952, 16, pp 90-129.
- (57) Rhine, J.B., and Pratt, J.G.: A review of the Pearce-Pratt Distance Series of E.S.P. Tests. J. Parapsychol. 1954, 18, pp 165-177.

- (58) Rivers, O.B.: An Exploratory Study of the Mental Health and Intelligence of E.S.P. Subjects. *J. Parapsychol.*, 1950, 14, pp 267-277.
- (59) Rush, J.H., and Jensen, A.: A Reciprocal Distance G.E.S.P. Test with Drawings. *J. Parapsychol.*, 1949, 13, pp 122-154.
- (60) Schmeidler, G.R.: Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *J. Amer. Soc. Psychic. Res.*, 1943, 37, pp 103-110.
- (61) Schmeidler, G.R.: Predicting good and bad scores in a clairvoyance experiment: A final report. *J. Amer. Soc. Psychic. Res.*, 1943, 37, pp 210-221.
- (62) Schmeidler, G.R.: Separating the sheep from the goats. *J. Amer. Soc. Psychic. Res.*, 1945, 39, pp 47-49.
- (63) Schmeidler, G.R.: Progress report on further sheep-goat series. *J. Amer. Soc. Psychic. Res.*, 1946, 40 pp 34-35.
- (64) Schmeidler, G.R., and Murphy, G.: The influence of belief and disbelief in E.S.P. upon individual scoring levels. *J. exp. Psychol.*, 1946, 36, pp 271-276.
- (65) Schmeidler, G.R.: Rorschach variables in relation to E.S.P. scores. *J. Amer. Soc. Psychic. Res.*, 1947, 41, pp 35-64.
- (66) Schmeidler, G.R. and Allison, L.W.: A repetition of Carington's Experiments with Free Drawing *J. Amer. Soc. Psychic. Res.*, 1948, XLII, pp 97-107.
- (67) Schmeidler, G.R.: Personality correlates of E.S.P. as shown by Rorschach Studies. *J. Parapsychol.*, 1949, 13, pp 25-31.
- (68) Schmeidler, G.R.: Some relations between picture-frustration ratings and E.S.P. scores. *J. Personality*, 1950, 18, pp 331-344.
- (69) Schmeidler, G.R.: E.S.P. Performance and the Rorschach Test. *J. Soc. psych. Res. Lond.*, 1950, 35, pp 323-339.
- (70) Schmeidler, G.R.: Personality Dynamics and Psychical Research. *Bull. Menniger Clin.*, 1954, 18, pp 19-30.
- (71) Scott, C.: Experimental Object-Reading: A Critical Review of the work of Dr. J. Hettinger. *Proc. Soc. psych. Res. Lond.*, 1949, 59, pp 16-50.
- (72) Sechehaye, M.: *Autobiography of a Schizophrenic Girl*. Grune & Stratton, New York, 1951. (P 41)
- (73) Shulman, R.: A Study of Card-Guessing in Psychotic Subjects. *J. Parapsychol.*, 1938, 2, pp 95-106.
- (74) Sidgwick, H, and others: Report on the Census of Hallucinations. *Proc. Soc. psych. Res. Lond.*, 1894, X p 33.
- (75) Sidgwick, E.M.: Psychology of Mrs. Piper's Trance Phenomena. *Proc. Soc. psych. Res. Lond.*, 1915, XXVIII, pp 305-314 & 622-624.

- (76) Smith, B.M., and Humphrey, B.M.: Some Personality characteristics related to E.S.P. performance. *J. Parapsychol.*, 1946, 10, pp 269-289.
- (77) Spencer Brown, G.: Statistical Significance in *Psychical Research*. *Nature, Lond.*, 1953, 172 p 154.
- (78) Soal, S.G., and Bateman, F.: Agents in opposition and conjunction. *J. Parapsychol.*, 1950, 14, pp 168-192.
- (79) Soal, S.G., and Pratt, J.G.: E.S.P. Performance and Target Sequence. *J. Parapsychol.* 1951, 15, pp 192-215.
- (80) Stuart, C.E.: The effect of rate of movement in card-matching tests of extra-sensory perception. *J. Parapsychol.*, 1938, 2, pp 171-183.
- (81) Stuart, C.E.: An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *J. Parapsychol.*, 1941, 5, pp 99-137.
- (82) Stuart, C.E.: An E.S.P. Test with Drawings. *J. Parapsychol.*, 1942, 6, pp 20-43.
- (83) Stuart, C.E.: The Carington Free-Drawing Approach to the E.S.P. Problem. *J. Parapsychol.*, 1944, 8, pp 127-138.
- (84) Stuart, C.E.: A classroom E.S.P. experiment with the Free-Response method. *J. Parapsychol.*, 1945, 9, pp 92-105.
- (85) Stuart, C.E.: An E.S.P. Experiment with Enclosed Drawings. *J. Parapsychol.*, 1945, 9, pp 278-295.
- (86) Stuart, C.E.: G.E.S.P. Experiments with the Free-Response Method. *J. Parapsychol.*, 1946, 10, pp 21-35.
- (87) Stuart, C.E.: An Interest Inventory Relation to E.S.P. Scores. *J. Parapsychol.*, 1946, 10, pp 154-161.
- (88) Stuart, C.E., Humphrey, B.M., Smith, B.M., and McMahan, E.: Personality measurements and E.S.P. tests with cards and drawings. *J. Parapsychol.*, 1947, 11, pp 118-146.
- (89) Symonds, P.M.: *Diagnosing Personality and Conduct*. D. Appleton-Century Co., New York, 1931.
- (90) Taves, E., Murphy, G, and Dale, L.A.: American Experiments on the Paranormal Cognition of Drawings. *J. Amer. Soc. Psychic. Res.*, 1945, 39, pp 144-150.
- (91) Taves, E.: The Construction of an American Catalogue. *J. Amer. Soc. Psychic. Res.*, 1945, 39, pp 151-156.
- (92) Thouless, R.H., and Wiesner, B.P.: The Psi Process in Normal and "Paranormal" Psychology. *Proc. Soc. psych. Res. Lond.*, 1947, XLVIII, pp 177-196.
- (93) Tyrrell, G.N.M.: *The Personality of Man*. Penguin Books, Harmondsworth, 1946.
- (94) Ullman, M.: On the nature of resistance to Psi Phenomena. *J. Amer. Soc. Psych. Res.*, 1952, XLVI, pp 11-13.
- (95) Warcollier, R.: *Mind to Mind*. Creative Age Press, New York, 1948.

- (96) Warner, L.: A second survey of psychological opinion on E.S.P. J. Parapsychol., 1952, 16, pp 284-295.
- (97) Washburne, J.N.: Washburne S-A Inventory, (Thaspic Edition) World Book Co., New York, 1940.
- (98) West, D.J.: Mass Experiments on the Psi-Cognition of Drawings. J. Soc.psych.Res. Lond., 1947, 34, pp 43-54.
- (99) West, D.J.: E.S.P. Performance and the Expansion-Compression Rating. J. Soc. psych. Res. Lond., 1950, 35. pp 295-308.
- (100) Woodruff, J.L., and Dale, L.A.: Subject and experimenter attitudes in relation to E.S.P. scoring. J. Amer. Soc. Psychic. Res., 1950, 44, pp 87-112.

Supplementary references:

- (101) Brownlee, K.A.: Industrial Experimentation, (Fourth Edition). H.M.S.O. London, 1951.
- (102) Fisher, R.A.: Statistical Methods for Research Workers. Cliver and Boyd, London, 1946.
- (103) Garrett, H.E.: Statistics in Psychology and Education, (Second Edition). Longmans, Green and Co., London, 1941.
- (104) Milton Smith, G.: A Simplified Guide to Statistics for Psychology and Education, (Revised and Enlarged). Rinehart and Company Inc., New York, 1949.
- (105) Moroney, M.J.: Facts from Figures. Penguin Books A236, Middlesex, 1951.
- (106) Rhine, J.B. and Pratt, J.G.: Parapsychology, Frontier Science of the Mind. Charles C. Thomas, Illinois, U.S.A. 1957.
- (107) Siegel, S.: Nonparametric Statistics for the Behavioural Sciences. Mc.Graw Hill Book Co., Inc., New York, 1956.