EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 3

NUMBER 1

NOVEMBER 1979

Robert L. Morris
Psychology Department
University of Edinburgh

EDITORIAL STAFF

Martin Johnson
Sybo A. Schouten

REGIONAL EDITORS

U.K.
Dr. John Beloff
University of Edinburgh
Department of Psychology
7 George Square
Edinburgh EH8 9JZ

W. Germany
Prof.Dr. Hans Bender
Prof.Dr. Johann Mischo
Institut für Grenzgebiete der
Psychologie und Psychohygiene
78 Freiburg i. Br.
Eichhalde 12

Sweden
Dr. Rolf Ejvegaard
Mellansjövägen 35
S-14146 Huddinge

Iceland
Dr. Erlendur Haraldsson
University of Iceland
Department of Psychology
Reykjavik

U.S.A.
Ms. Rhea White
2 Plane Tree Lane
Dix Hills
New York 11746

Italy
Dr. P. Cassoli
39, Via Valeria i
Bologna 12

Australia
Dr. H.H.J. Keil
University of Tasmania
Box 252 C, G.P.O.
Hobart, Tasmania

Prof.Dr. Emilio Servadio
Via di Villa Emiliani, 4-int.7
I-00197 Roma

The European Journal of Parapsychology publishes original experimental and theoretical papers on parapsychology. In addition, articles presenting reviews of literature, case studies and papers on related topics are accepted in so far as the subject can be considered relevant to the understanding of parapsychological phenomena or to the methodology of empirical research in parapsychology. Translations of papers originally published in a language other than English are also welcome.

Manuscripts should be written in English and should be typed, double-spaced throughout, with a wide margin, on one side of the page only. The organization of the paper must be clearly indicated by appropriate headings and subheadings. Include names of authors, positions and other affiliations, and mailing addresses. Each table should be given a descriptive title. Number tables consecutively with Arabic numbers in order of their mention in the text. Tables should be designed to fit into a space of 12 x 19 cm. or less.

(continued inside back cover)

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

CONTENTS

Johnson, M.U.	The European regional branch of the Parapsychological Association (ERPA) is now established	p. 1
Krippner, S.	Gardner Murphy: some memories	p. 3
Haraldsson, E. Johnson, M.U.	ESP and the defense mechanism test (DMT) Icelandic study no. III. A case of experimenter effect?	p. 11
Keil, J.	Field experiments with 30 possible PK subjects	p. 21
Gruber, E.R.	Conformance behavior involving animal and human subjects	p. 36
Pasricha, S Stevenson, I.	A partly independent replication of investigations of cases suggestive of reincarnation	p. 51
Johnson, M.U.	Obituary J. Gaither Pratt	p. 66
Beloff, J.	The categories of psi: the case for retention	p. 69
Millar, B.	The distribution of psi	p. 78
	Letters to the editor	p.111
	Erratum	p.117

PUBLICATION POLICY

Twice a year the Parapsychology Laboratory of the University of Utrecht publishes the European Journal of Parapsychology. The object of the European Journal of Parapsychology is to stimulate and enhance the activity in this field, especially in our corner of the world, by communicating research results and issues related to professional parapsychology. Although there will be an emphasis on experimental work, theoretical articles are also welcome. Contributions from all over the world will appear in the journal.

A hallmark of the European Journal of Parapsychology is the attempt to avoid selective reporting, that is, the tendency to bury 'negative' results and only to publish studies that 'turn out'. To avoid turning the journal into a graveyard for all 'unsuccessful' studies, we require that the acceptance or rejection of a manuscript should take place prior to the phase when the experimental data are collected. The quality of the design and methodology and the rationale of the study are considered more important than the level of significance of the outcome of the study. As a practical rule, we advice the potential contributor of an article to submit the design of his planned study before the work is actually carried out. The rationale of the study should be stated, as well as all the hypotheses related to it. Furthermore one should try to specify the number of subjects, the number of trials, etc., plus the type of statistical methods one plans to use for evaluation.

Priority will be given to the publication of studies which fulfil the above-stated publication policy.

The final manuscript with presentation of results must reach us two months in advance of the official publication dates, which are May 1st and November 1st.

THE EUROPEAN REGIONAL BRANCH OF THE PARAPSYCHOLOGICAL ASSOCIATION (ERPA) IS NOW ESTABLISHED

Martin Johnson

The outcome of a survey as to the interest among European Full Members and Members of the Parapsychological Association of establishing the ERPA turned out to be an almost unanimous acceptance of the idea, in spite of the fact that it implies a slightly increased fee.

The return rate of the survey was 65 per cent (in spite of the fact that it was made clear that if we did not receive an answer we would interpret that as a sign that the person sympathized with the proposal).

Only four per cent of the answers indicated that they did not entirely sympathize with the idea.

In view of the massive acceptance of the idea of the establishment of the ERPA, the Council of the P.A. decided at one of its meetings at the 22nd Annual Convention at Moraga, California,U.S.A.,to endorse the establishment of the ERPA. It was decided that the new fee for Europeans (which now equals that of American Full Members/Members) has to be paid from 1980 onwards. The earmarked part of the fee ($ 5) will subsequently be transferred to one of the ERPA Committee representatives. A working-committee for the next two years was established at the P.A. Council meeting. Besides the undersigned, the following persons were elected:

Dr. John Beloff, Edinburgh
Prof. Dr. Hans Bender, Freiburg
Dr. Erlendur Haraldsson, Reykjavik

During this period a constitution for the ERPA will be drawn up for circulation and final approval.

It might be useful to iterate the main objectives of the ERPA as we see it:

1) To arrange regional colloquia or review-meetings at which Full Members and Members of the P.A., actively engaged in

research can discuss research projects under way, or research problems leading to the outline of an experiment. The meetings would offer an excellent opportunity for exchange of ideas and for obtaining useful criticism before the experiments have (actually) been carried out.

1.1) The meetings should also be helpful in integrating disparate research efforts. One of the functions of the meetings could be that they might serve as a clearing-house regarding the kind of contrivances available in the field - for instance, the kind of hardware and software that exist at a certain research center.

2) A second and no less important function of a European branch should be to stimulate and guide students and persons interested in carrying out some piece of research but living in areas which are less integrated in the main stream of parapsychological research. It should give them a chance to see and learn and to obtain many pieces of information of importance for their design and the carrying out of a field study or laboratory work. In this way the organization could be of great educational importance and could exert a guiding or steering influence on enthusiastic but less sophisticated newcomers to the field. These may need help and supervision to be able to carry out a study which ideally could lead to a paper of such quality that it could be considered for presentation at the regular P.A. Convention.

It seems appropriate to suggest that next year's S.P.R. International Conference, to be held in April, 1980, would constitute a good opportunity for arranging the first ERPA-meeting, where different issues pertinent to the stated objectives could be discussed.

It is imperative that Members in different corners of Europe try to identify and to encourage potentially able and dedicated research-workers to come and take part in such a meeting, where they would have a chance to capitalize on criticism of their research-projects before actually carrying out the intended research.

Should there be need for additional information, please, do not hesitate to write me or to contact any other member of the working-committee whom you think could be helpful in rendering you further information.

GARDNER MURPHY: SOME MEMORIES

Stanley Krippner

Humanistic Psychology Institute

I first met Gardner Murphy when he gave a guest lecture at Northwestern University's psychology department. I had read several of his books and had heard him speak at the annual convention of the American Psychological Association. However, it was not until 1959 that I had the pleasure of spending some time with him. A small group of students interested in parapsychology had prevailed upon him to meet with us after his departmental lecture was concluded.

It worked out that we were able to drive GM (as Gardner Murphy was frequently referred to) to Chicago to catch his train to Kansas, where he worked as director of research for the Menninger Foundation. En route, Murphy remarked that he was looking forward to spending the summer as a visiting professor at the University of Hawaii. I replied that I had been invited to Honolulu to participate in a friend's wedding; GM suggested that I become his teaching assistant.

Social psychology in Honolulu

The summer of 1960 was everything I could have hoped for. Each morning, I assisted GM in his undergraduate class on social psychology as well as his graduate seminar. At noon, Lois Murphy would join us for lunch where we discussed parapsychology, personality theory, and international affairs. One day we received word that a group of townspeople had discovered GM's presence on campus and threatened to picket his classes because of his purported "Communist" leanings. Their ammunition was scanty, consisting of GM's (1953) direction of a UNESCO project in India and his chairing (and co-founding) of the Society for the Psychological Study of Social Issues. Further, a tract from the arch-conservative publication Current Events had been circulated in which E.J. Janisch (1959) had accused GM of "playing the Soviet's game".

> ...Murphy, brilliant author of an important work on the theory of personality, spoke to a large college convocation this past summer. His topic? Understanding. Unless we begin

> to understand one another, we will blow each other up. On all the "liberal" juke-boxes, the number one hit is the same: Understanding .... There is plentiful proof that any benefits from understanding are .... highly speculative, and outright dangerous if entered into uncritically. I suggested all of this to Dr. Gardner Murphy, and he practically told me that I was an old mossback ...

GM's eyes twinkled as he discussed the forthcoming excitement; I believe that he was disappointed when the picketers failed to arrive. A few weeks later, GM asked me to buy him a Russian-English dictionary; he and Lois planned to leave for Moscow immediately after the final exams had been graded. I jested that the picketing would take place after all, once his trip to the USSR was revealed.

By the end of the summer, I had friends among the Japanese, Chinese, Filipino and Hawaiian ethnic groups. They enjoyed taking me to different parts of the island for new experiences with music, food, and entertainment. But the day before Murphy was to leave, I dined alone in a restaurant serving mainstream American food. I awakened in the middle of the night; I was nauseous, and after my third trip to the bathroom, concluded I had food poisoning.

It took me an hour to walk the few blocks to the campus, carrying the exam questions with me. I explained to the class I was ill and that they should not mind me as I made my frequent excursions to the bathroom. In the meantime, I graded the essays as best I could and delivered them to GM a few hours later. He was shocked to see my condition and remarked at the green tone my face had assumed. He wanted to take me to a hospital but I insisted that we decide on final grades. This done, I told him to head for the airport or he would miss his flight. He agreed but said he would feel better if I were in the hospital.

With GM on his way to meet Lois for their overseas trip, I delivered the grades to the registrar's office, then went to bed. However, I had no intention of going to the hospital; I had planned to go mountain climbing the next day and did not want to miss the experience.

A decade of dreams

A year later I graduated from Northwestern University and

became director of the Child Study Center at Kent State University. A friend and I visited the Murphys at the Menninger Foundation where we discussed the research projects he was supervising. We also met yearly at the meetings of the American Psychological Association and the Parapsychological Association. At the 1963 PA convention, GM introduced to me Montague Ullman and in 1964 I moved to Brooklyn to investigate ESP in dreams at the Maimonides Medical Center Dream Laboratory.

During the ten years that I directed the laboratory's research activities, GM was a frequent guest. His trips to Brooklyn were especially popular with our student assistants because of the personal interest he showed in them. I had observed this trait many times, one of the first being at an APA meeting when he was the guest at a "conversation" hour. Instead of delivering a monologue, as did most of the guests I had observed on these occasions, GM actually conversed with the roomful of people who had come to meet him. He asked them what ideas they had found stimulating and what they had enjoyed at the convention. He would then share his opinions and feelings on the topic at hand. I watched as people beamed, basking in the knowledge that GM thought that their experiences were worth sharing.

GM's eclectic approach to knowledge generated a high degree of creativity. While at Menninger's, his concepts produced some of the first data demonstrating that individuals had a remarkable -- and hitherto unsuspected -- ability to control their own physiology. GM was familiar with yogic psychology and with the human autonomic nervous system; in his landmark book, Personality, he laid the basis for biofeedback research. In 1969, an organizational meeting was held of clinicians and researchers interested in studying the voluntary control of internal states. Although the group eventually became known as the Biofeedback Society of America, an initial proposal suggested they adopt "The Gardner Murphy Society" as their name.

Earlier that same year, I had been invited by the Menninger Foundation to attend the First Interdisciplinary Conference on the Voluntary Control of Internal States. GM was present although he had joined the psychology department of George Washington University as visiting professor following his retirement from the Menninger Foundation in 1967. The event must have been gratifying to GM; almost every presentation covered material in which he had been a pioneer -- although he was too modest to point this out.

Murphy's modesty had been apparent to me ever since our days

together at the University of Hawaii. I had expected him to use one of his books as a required text for the social psychology course; instead, he selected two books by other authors. However, he was not modest in a self-deprecating way. Although I had heard him express his eclecticism by stating, "I am a Freudian, a Jungian, an Adlerian, a Rogerian," he also said, "Why should there be a Murphian psychology? There is no Darwinian biology."

The creative life

In May 1968, the National Association for Gifted Children held its fifteenth annual convention in Chicago. I was selected to give the organization's annual Citation of Honor to Gardner and Lois Murphy. G.M. arrived an hour before the banquet and has a pleasant chat about our current research projects at Maimonides before joining the others. In presenting the award, I noted that GM had received his degrees form Yale, Harvard, and Columbia, and how he had served as president of the American Psychological Association, the Society for Psychical Research (in Great Britain), and the American Society for Psychical Research. In accepting the award, GM made six cogent observations:

1. We are wasting thousands of new lives through grossly inadequate mother and infant health care.
2. We are wasting many potentially fine minds through a static view of the intelligence quotient, instead of grasping the ...possibilities latent in the complex interaction of genetic and opportunity factors.
3. We are paralyzing a large portion of our teachers by rigid demands and limitations upon them ...and by sticking to stereotype rules rather than encouraging teachers to develop their resources.
4. We are stultifying a range of potentialities of young minds by exposing them to the repetitive cliches and violence-oriented television programs instead of utilizing these resources for the stimulation of creative imagination.
5. If only we could liberate the potential abilities of the thousands of youths who become dissenters, protestors, hippies and children of despair, or had the guts to demand a saner and more loving society, we could make constructive use of the energies underlying these protests.
6. We need courage to develop the creative vision of human goals on the part of dedicated teachers, doctors, social workers, nurses, and all the helping groups which can point the way, each one offering his or her creative help as a gift to all children, however handicapped or

> however brilliantly gifted. ...If we once begin to think in individual terms, individual interactions between adults and children, our goals will become less abstract, more personal, more capable of bringing to creative life whatever lies latent within each individual.

Hearing these words, I recalled what I knew about GM's own childhood. The son of an Episcopalian minister and a Vassar-educated teacher, GM was born in Chillicothe, Ohio, on July 8th, 1895. The family moved to New England where his mother taught him about birds as a child and his father, the author of an introductory astronomy book, taught him about stars.

GM often gave credit to his wife for stimulating his interest in children, and he once wrote that his personality concepts "were derived in large measure from the rich empirical material on child development which she was constantly working through." Their brilliant book on experimental social psychology (Murphy and Murphy, 1931) was a landmark in the field. GM stated, "It is self-evident to anyone who knows me that I have learned a great deal more psychology from Lois than from any other living person" . It was also self-evident to anyone who knew them that theirs was one of the most fruitful working relationships in psychology as well as one of the great love stories of the age.

As a child, GM was interested in active sports such as baseball and football. During World War I, he served in the American Expeditionary Forces and the Yale Medical Corps. But in 1925 he became seriously ill, having succumbed to the grippe. He was a semi-invalid for many years and during much of this time he was unable to read. In 1934, having tried everything recommended by orthodox medicine, he heard of Dr. W.H. Hay's regimen of diet and physical exercise, regarded by the medical establishment as sheer quackery. However, GM's health was restored following a month on the program. This experience was instrumental in keeping his mind open in regard to the unorthodox.

GM's limitless range

In 1972, I visited the USSR and India, recalling GM's previous trips as I walked by the Kremlin and, later, observed the shores of the Ganges River from a boat. On my way back to New York, I stopped in Honolulu for the annual convention

of the American Psychological Association. As usual, I had lunch with the Murphys, but this occasion was marked by a pleasant surprise. Lois confided that later that week GM would receive the Gold Medal Award, his profession's highest honor. I was in the audience when the award was presented to GM who was described as "a peerless teacher, a felicitous writer, an eclectic psychologist of limitless range, he seeks to bring the whole of human experience to bear in understanding behavior." Later, we reminisced about our summer in Hawaii -- and I confessed that I had gone mountain climbing rather than to the hospital following my bout with food poisoning.

GM was even willing to give astrology a chance to demonstrate its validity. In 1972, he was approached by an astrologer who wanted to take part in a scientific evaluation of his craft. I obtained 16 volunteers who gave me their birth information. The astrologer constructed horoscopes of each of the subjects. However, I had taken care to insure that half the subjects were almost completely naive about astrology while the other half knew their "sun sign, "moon sign", "rising sign", and the characteristics of their "houses".

We gave the members of this latter group eight personality descriptions -- those written out by the astrologer (with any references to specific "signs" deleted). We gave the members of the naive group the eight personality descriptions corresponding to them. Each group rated (on a 100 point scale) the similarity of each personality description to their self-concept. A statistician evaluated the results. Neither group obtained a significant result. One member of the naive group correctly identified his personality description (giving it the highest score of the eight) but this is what one would expect from chance. Two members of the sophisticated group made correct identifications, and the others tended to give higher scores to the correct descriptions than did members of the naive group. The average score given to their horoscopes by members of the naive group was 48; they gave the other horoscopes an average score of 44. The sophisticated group did better, giving average scores of 68 and 56. These results were not statistically significant. Further, those in the sophisticated group might have zeroed in on key words they knew to be associated with their signs (such as a "headstrong" and "determined" Aries), accounting for the more favorable scores given their own horoscopes. We did not claim to have conducted a definitive study of astrology; however, we had constructed a useful design which could be used in future investigations.

Celestial orbits

In 1970, GM had given the keynote address at a parapsychology conference at the University of Maine. I recall hearing him stress the need for replicable ESP experiments, and praising a study we had done at our dream laboratory, obtaining similar results in a replication study with the same subject, same telepathic "sender", same experimenters, and same evaluators (Ullman, Krippner, and Vaughan, 1973, Chapter 9). On virtually every visit, he would query me as to what replications had occurred in the field.

Our conversations were a joy and a delight, but I did miss his letters once he could no longer write them, due to illness. When the Special Collections Department of the Kent State University Library asked me for my personal papers, I confessed that I did not keep copies of my correspondence, not thinking it of any historical value. However, I did keep files on letters that people like GM had sent to me. The originals of GM's 62 letters (and those of Charlotte Bühler, J.B. Rhine, Alan Watts, etc.) were taken by courier to Ohio -- as well as the lecture notes that GM had written out for his social psychology course in Hawaii during 1960.

However, there was one letter each year that continued to arrive -- the "happy holidays" newsletter, often containing a poem which GM had dictated. And during each of my visits, his good humor and high spirits were evident. One visit coincided with a writing-session GM was holding with Morton Leeds. GM remarked how we all were positioned at different points in the room: "If I were the sun, Mort would be Mercury, Stan would be Venus, Lois would be Earth, and the nurse, Mars." We all had a chuckle as we continued our celestial orbits!

I last saw GM in January, 1979. On March 18, GM had a conference with Morton Leeds. Lois entered the room and GM held out his arms to embrace her. He was helped to get into his walker for some exercise and spoke of his desire to use a cane during the annual summer trip he and Lois would make to their beloved New Hampshire retreat house. He went back to his favorite chair in time to watch a ballet performance on television. Then his heart stopped beating.

When I received the news, I was saddened. But I also knew that GM's last years represented a triumph of the human spirit over illness. His physical condition was shocking to many. An 18-inch gastrostomy tube had been inserted into a hole in his

stomach through which he was fed by carefully pouring purees and liquids. This was necessary because Parkinson's disease had made swallowing difficult, causing food to become aspirated into his lungs. A tracheostomy (i.e., a tube inserted into a hole in his throat) permitted the elimination of secretions when he could not cough during a cold. Both of these procedures caused recurrent pain, but when the pain was relieved with topical anesthetics and ointments, GM seemed unaware of these intrusive pieces of apparatus. Parkinsonism also caused almost continuous tremors of his hands, which turned his once firm handshake into a quivering greeting. The disease made speech difficult and voluntary adjustments of his body almost impossible. However, these frustrations did not affect his wit, his intellectual vitality, or the twinkle in his eye. He finished two of his most engaging books during this period -- Outgrowing Self-Deception and The Paranormal and the Normal.

In September, 1979, a special memorial meeting was held at the American Psychological Association convention. For once, I could not be present because the meeting was in direct conflict with a presentation of mine. However, my friends told me that GM's parapsychological interests were applauded by some speakers, treated with reserve by others, and ignored by a few. None, however, doubted the greatness of the man or his innate goodness. To many of us GM will also stand for "Great Man".

REFERENCES

Murphy, G. & Murphy, L.B. Experimental Social Psychology. New York: Harper & Row, 1931.

Ullman, M., Krippner, S. & Vaughan, A. Dream Telepathy. New York: Mac Millan, 1973.

Stanley Krippner
Humanistic Psychology Institute
San Francisco, California
U.S.A.

ESP AND THE DEFENSE MECHANISM TEST (DMT)
ICELANDIC STUDY NO. III.
A CASE OF EXPERIMENTER EFFECT?

Erlendur Haraldsson
University of Iceland

Martin Johnson
University of Utrecht

In much current work it is assumed that psi-missing is the result of some psychological mechanism of distortion or defense. The Defense Mechanism Test (Kragh and Smith 1970) is said to measure psychological defenses of some kind. For theoretical reasons therefore research into the relationship between psi and the Defense Mechanism Test (DMT) has been seen as of considerable importance: such work has promised to shed light on those murky waters in which psi seems to operate. Prior to our Icelandic investigation the results of five experiments had been published and all of them showed a significant relationship between ESP and the DMT. In these experiments correlations ranged from .26 to .79. These results reveal a rather remarkable degree of repeatability in our field, all the more so for the important reason that we know with certainty of all studies that have been conducted since only a handful of psychologists know how to score the DMT. Hence we can be sure that there is no queue of unascertainable length consisting of unsignificant unpublished experiments.

In view of the previous successes of the test as a predictor in a parapsychological setting it was considered important to conduct a series of experiments to explore further the usefulness of the DMT as a predictor of ESP performance. We have now conducted three experiments: one has already been published (Haraldsson 1978); one, which gave the first non-significant correlation found, has been reported at a recent conference of the Society for Psychical Research; in this paper we report on a further experiment. Before discussing the present investigation a description will be given of the DMT itself since few are acquainted with the test.

Note: This publication fulfils the publication policy of this journal.

The background of the DMT

The DMT was invented and to a large extent developed by the Swedish psychologist Ulf Kragh, currently professor of applied psychology at Lund University.

The test can briefly be described as a projective test in which TAT-like pictures are presented tachistoscopically. The stimuli, administered by the use of a tachistoscope are utilized to measure what has sometimes been referred to as "perceptual defense", or according to Professor Kragh's terminology "preconscious defensive organization", (PDO), in a person's perception.

The rationale of the DMT is derived from a) psycho-analytic assumptions about the function and dynamics of defense mechanisms; b) from certain general concepts related to projective techniques such as the Thematic Apperception Test, the TAT; c) from the German school of psychology known as Actual Genese, in which the relevance of the developmental aspect of perception is emphasized. In the paradigm of Percept Genetic Analysis, mainly developed by Professor Kragh and Professor Smith at Lund University, much emphasis is also put on the developmental and sequential analysis of perception. The tachistoscopic technique makes it possible to perform a detailed analysis of how the perception of certain stimuli built up. Finally, concepts related to "subliminal perception" and in general to preconscious processes are of central importance to the DMT-technique.

A few words should be said about the psychoanalytic key-idea behind the DMT. According to S. Freud's scheme for the dynamics underlying defense mechanisms the following three links are relevant.

I. Threat/Danger -- Anxiety -- Defense against Anxiety

In an experimental setting Kragh has transformed this scheme into a dynamic triad of structures:

II. Subliminal threat in Perception __ Submanifest Anxiety __ Defensive Responses in Perception

In the DMT subliminal threat is induced by short-term exposures of TAT-like, threatening pictures.

About the DMT-pictures

Slides are used on which a centrally located person is drawn. This central person is referred to as the hero person. Somewhat peripherally and somewhat obscured and standing behind

the hero a very unpleasant looking "secondary person" is drawn, stretching out to reach the hero. This "secondary person" is referred to as the "threatening person". If male subjects are tested, slides are used which depict a young male person and a threatening person (usually also male). At full-time exposure the threatening person is judged to be approximately 20 years older than the hero person.

The administration of the DMT

A rather detailed instruction is read to the subject or subjects. The test can be administered both to individuals and to groups. There are noticeable advantages associated with individual administration of the test, but one shortcoming is that the procedure takes much more time. For individual testing a tachistoscope is used which consists of a viewer and an electronic timer. For group testing, a tachistoscope is used which consists of a slide projector to which a camera shutter is attached. By the use of the camera shutter, the exposure times can be varied from very short exposures up to rather long ones. In the group-testing situation the pictures are flashed onto a transparent screen of defined size. The projector and the subjects are located on different sides of the screen at defined distances. The illumination level of the room and of the tachistoscope are controlled. In the group-testing situation the number of subjects is usually restricted to 5 or 6. Two "parallel series" (e.g. pictures) are administered in order to increase the reliability of the test.

Exposure times

The duration of exposures in an individual testing increases logarithmically from 10 to 2000 milliseconds (via 20 steps). In a group session the exposure times usually range from 8 milliseconds (a sufficiently short exposure time to be well below the physiological threshhold of any subject) up to 500 milliseconds.

The coding of information

After each exposure a subject has to report what he thought he saw, both by a written report and by making drawings and so-calles "markings". These "markings" refer to whether the S thought he saw one or more persons in the picture; of what sex he thought the person or persons were;

whether the person or persons were sympathetic or unsympathetic etc. Any change from one exposure to the other should be reported meticulously. This allows for a detailed "serial analysis" of what a subject experiences, or at any rate is willing to report.

The evaluation of a DMT-protocol

The evaluation procedure is a somewhat complicated two-step one. It takes several months of intensive training to become a reliable DMT-interpreter (the second step) whereas three to four weeks of intensive training may suffice for learning how to code (step one) the information given after each exposure.

Step one: A phenomenological analysis is carried out in which characteristics or "anomalies" in the report given after each exposure leads to the application of one or more codings (- recording of "signs" or "sign variants").

No such codings are made until the exposure where the S for the first time reports seeing a living being or an object at the place of the hero or the threat figure. Coded "signs" are thought to reveal that preconscious defense mechanisms have been activated in the subject's perception. These alleged preconscious defense mechanisms or PDOs ("preconscious defensive organization") are thought to be related to classical "mechanisms" such as repression, isolation, denial, reaction-formation and so forth, a view which is supported by empirical studies carried out on psychiatric patients.

Step two: Here a global assessment of the coded protocol is executed, a process in which consideration is given to types of "signs", mixtures of "signs", sequential aspects of the "signs" and correspondences of the "signs" in the two DMT-series. The final rating of the protocol is done based on data from validation studies and on theoretical considerations. For most purposes these ratings are expressed in a nine-(as in the current experiment) or five-point scale. A numerically low rating indicates that the subject is considered to be rather anxiety prone, which implies that S should have limited possibilities of successfully coping with emotional stress. Certain types of DMT-protocols given a low rating, are thought to indicate that the subject has great difficulties in

handling preconscious processes in general. It would also imply poor ability to recall dreams. High numerical ratings are thought on the one hand to indicate high ability to function well under emotionally stressed conditions; on the other to possess good ability to come to terms with preconscious processes. It would also go together with good ability to recall dreams.

The application of the DMT

The DMT was primarily developed as an instrument for selection within the armed forces, especially for prediction of a person's ability to cope with stressors implying a real threat towards his life. Therefore, the application and validation of the test has been closely related to the performance of Air Force pilots, parachutists and attack divers, categories which in their training and work are actually exposed to such stressors.

The standardization of the DMT has been carried out during the last 20 years. This work has enjoyed financial support from governmental sources in Sweden. Recently a Ph.D.-thesis was defended at Lund University, which deals with the DMT as a predictor of flying performance within the Royal Swedish Air Force. (Neuman, 1978.) In this thesis it was shown that:

a) Accident-proneness among pilots can be predicted rather well by the DMT;
b) It was also shown that the DMT has more predictive power left after the pass/fail criterion has been applied for the training of pilots, than any other of the psychological instruments used.

One should also note that the DMT has a quite low correlation with other paper and pencil-tests, including questionnaires supposed to measure manifest anxiety. Today, the DMT is used for selection of pilots in all the Scandinavian Air Forces.

A person's ability to recall dreams may also be fairly well predicted on the basis of his DMT-data. (Johnson, 1970.)

PROCEDURE

The procedure we have followed in the three Icelandic studies are in most respects identical. For Icelandic Study I, see Haraldsson, 1978.

In the present study (Icelandic Study III) each subject had to participate in two sessions, one for the DMT-testing and another, in the following week, for the ESP part of the experiment.

SUBJECTS

41 subjects, all male students and volunteers, took part in the experiment. They were told they were to participate in a perceptual test and in an ESP test.

To prevent knowledge of the nature of the DMT stimuli from spreading among prospective subjects, they were recruited from six different departments of the university, all subjects (6-8) from the same department participating in the DMT test at the same time. Eight students of psychology served as experimenters in the ESP part of the experiment and they also recruited the subjects. These experimenters were not completely naive about the nature of the DMT, but had only vague ideas about the rationale of the test. However, the method of evaluation was not known to them and during their participation in the experiment they never came into contact with the DMT-protocols.

The DMT-testing

Most of the planning and preparation of the experiment was done by Erlendur Haraldsson. Martin Johnson came to Iceland to carry out the DMT-testing. Many of the Ss wrote their protocols in English, but protocols in Icelandic were verbally translated to Martin Johnson by Erlendur Haraldsson after the DMT-sessions and during Martin Johnson's five days stay in Iceland.

Rating and final evaluation of the DMT were carried out by Martin Johnson after he had left Iceland. The ESP testing also took place after Johnson's departure.

The ESP task

As in the previous two experiments the ESP task consisted of 80 trials, 40 of clairvoyance and 40 of precognition. An electronic machine with four push-buttons (an "ESP-tester", produced by Paratronics, Las Gatos, California) was used for the clairvoyance test. For each trial the subject indicated his choice of a correct target by pressing one of the buttons ($p = 1/4$). Two subjects were tested together and two attending

experimenters tried to develop a light and competitive spirit between each pair. To further increase motivation, every subject who obtained 15 or more hits in the clairvoyance test (MCE = 10) was offered a long-playing record of his own choice.

For a more detailed description of this task see Haraldsson 1978.

Each pair of experimenters tested only 4 to 6 pairs of Ss which they had also brought to the DMT session.

The 40 trial paper and pencil precognition test was conducted immediately after the clairvoyance test. For each trial the subject had to make a cross in one of four boxes that a noise driven RNG/computer system at the State University of Utrecht would randomly select for each subject a few days after all the subjects had completed the experiment. Sybo A. Schouten was responsible for generating the precognition targets.

On a predetermined date after all ESP sessions were over Erlendur Haraldsson mailed the ESP data to Utrecht and Martin Johnson mailed the DMT data, and Sybo Schouten the precognition targets, to Reykjavik. Thus, an independent treatment of the ESP and DMT data was ensured.

RESULTS

When the total ESP scores (for clairvoyance and precognition combined) were correlated with the DMT ratings we obtained a tiny and non-significant correlation coefficient of r_S = .02. For clairvoyance alone the correlation was .07 and for precognition .00. A further breakdown of these data can be seen in table 1.

Looking at the ESP scores we find for clairvoyance a mean of 9.10 (MCE = 10.00) which is significantly below chance (CR = 2.11) whereas the mean for precognition is 9.98, or close to mean chance expectation.

As 4 pairs of experimenters conducted the ESP part of the experiment, we analysed their effect on the ESP-scoring by analysis of variance. No significant effect of these pairs of experimenters was found ($F = 1.96$, df 3/37, $P = .14$).

At the University of Iceland we have now conducted 3 experiments testing the ESP-DMT relationship. These

Table 1

Means and standard deviations of total ESP performance (clairvoyance and precognition combined) in relation to DMT-ratings.

DMT-rating	Number of Ss	Mean ESP score	St. dev.
1 high defensiveness	2	18.50	2.12
2	3	21.00	1.73
3	6	19.50	2.74
4	6	18.00	3.03
5	8	17.25	4.80
6	6	19.83	5.71
7	5	20.40	2.88
8	3	19.67	3.21
9 low defensiveness	2	19.50	3.54
Total	41	19.07	3.69

experiments have been basically of the same design with only slight changes between experiments. In the first experiment we obtained a sizeable Spearman correlation of .46, in the second experiment an insignificant correlation of .17 and now we are down to .02.

It is interesting to look back at the history of ESP research with the DMT. As can be seen in table 2 the first experiment (Carpenter 1965) yielded a high correlation of .79. Since then correlation decreased with every experiment until it reached .26 (Johnson and Lübke 1977). That was just before the senior author of this paper obtained Martin Johnson's cooperation in conducting the first Icelandic experiment. Then the correlation went up again to .46 to drop in the third Icelandic experiment to an all-time low of .02.

The decline from .79 to .02 is impressive in its regularity and raises several questions. However a negative correlation has not been obtained in any of the eight experiments conducted so far and only two have not been significant. This alone reveals a rather remarkable degree of repeatability in this field. Are these results only fluctuations or a small but genuine positive correlation? Or are the significant positive correlations solely caused by the experimenters, presumably psi - mediated, but revealing nothing about an intrinsic relationship between ESP and DMT? Or are both factors perhaps involved in the results obtained so far?

Martin Johnson initiated all the first five experiments. He was the driving force behind them. Results of those experiments declined slowly from first to last. The three experiments conducted in Iceland were initiated by Erlendur Haraldsson and they also show a steady and steep decline. As a whole the results of ESP research with the DMT suggest an experimenter effect or an initial effect of some kind. But an important question remains: will the DMT-ESP relationship hold to some extent in further experiments or will it sink into oblivion as have some other promising correlates or predictors of ESP performance?

Table 2

Survey of experiments involving DMT and ESP

	No of Ss	r_s	P
Carpenter, 1965	10	.79	.01
Johnson & Kanthamani, 1967 I	16	.67	.005
II	11	.59	.05
Johnson, 1975	18	.42	.05
Johnson & Lübke, 1977	49	.26	.05
Haraldsson, 1978	37	.46	.002
Johnson & Haraldsson, 1979	38	.17	n.s.
Haraldsson & Johnson, 1979	41	.02	n.s.

REFERENCES

Carpenter, J.C. An exploratory test of ESP in relation to anxiety proneness. In Rhine, J.B. and Associates: Parapsychology from Duke to FRNM, Durham Parapsychology Press, 1965.

Haraldsson, E. ESP and the Defense Mechanism Test (DMT). A Further Validation, European Journal of Parapsychology, 2, 1978, 104-114.

Johnson, M. Dream report related to percept-genetic defensive organization in the DMT. In Percept-Genetic Analysis, edited by Kragh, U. and Smith, G. Gleerups, Lund, 1970.

Johnson, M. ESP and subliminality. European Journal of Parapsychology, 1, 1975, 9-17.

Johnson, M. and Haraldsson, E. An Unsuccessful Replication of a DMT-ESP Experiment. Paper given at the 3rd International Conference of the S.P.R., Edinburgh, April 1979.

Johnson, M. and Kanthamani, B.K. The defense mechanism test as a predictor of ESP scoring direction. J. Parapsychology, 31,1967, 99-110.

Johnson, M. and Lübke, C. A further attempt to validate the DMT as a predictor of scoring direction. European Journal of Parapsychology 1, 1977, 37-46.

Kragh, U. and Smith, G. Percept-Genetic Analysis. Lund, 1970.

Miller, S.W. and York, M.S. Perceptual defensiveness as a performance indicator on a free-response test of clairvoyance. In: Research in Parapsychology, 1975, 162-166, Metuchen, N.J., Scarecrow Press, 1976.

Neuman, T., Dimensionering och validering av percept-genesens försvarsmekanismer. En hierarkisk analys mot pilotens stressbeteende. FOA rapport C 55020-HG, Stockholm, October 1978.

FIELD EXPERIMENTS WITH 30 POSSIBLE PK SUBJECTS

Jürgen Keil

University of Utrecht

Psychokinesis of a kind which can be directly observed has been reported in recent years to a far greater extent than was the case only 10 years ago (e.g. see RIP and Journal ASPR). PK has also started to attract more theoretical considerations, so much so that arguments are advanced according to which PK is the basis of psi events and ESP is only seen as a special case of PK (Millar, 1978). Future experimentation with PK is likely to provide important clues to the understanding of psi.

In the past directly observable PK was associated only with a handful of rare individuals and the disadvantages resulting from this situation may at times have outweighted the advantage of finding answers more quickly and of reducing the control problems mainly to the avoidance of trickery and/or unintentional manipulations.

It is of considerable interest to ask what has brought about the increased PK occurrences and how to maintain and if possible to increase the number of directly observable PK investigations.

Most likely there are large scale factors in operation which have a bearing on the questions as to how many people will consciously experience PK and how many will accept it as reality and how many will talk about it and possibly report it to parapsychologists. In the outlook of the general public there have been changes towards the acceptance of a less rigidly determined universe. The mass media played an important role in pointing out differences between scientists discussing orthodox fields as well as reporting claims and counterclaims in connection with parapsychology. Changes in the religious outlook such as the acceptance of charismatic subgroups among the more orthodox Christian denominations and the increasing interest in the teachings of non Christian religions probably also provided the necessary conditions for some people. The higher awareness of uncertainties on account of technological changes and of crisis situations - often reported by the mass media out of proportion to other events - may also have contributed to a general change

towards a less rigid outlook. Nevertheless, it has to be recognized that it was not until Geller's public performances reached thousands of viewers when many people were triggered into experiencing what they believed to be paranormal events. On account of more recent changes in the general attitude of the public it may well be that Geller is now significantly less able to obtain the kind of public response which he was able to gain in 1974. It must be noted though that even in 1974 there were substantial differences between countries.

Parapsychologists can only influence the general attitude of the public to a small extent through reasonable educational programmes. It is of some importance to investigate specific factors which seemed to operate in connection with Geller and which inspired many people to attempt PK and which probably allowed more to succeed with it than was ever achieved with the best experimental programmes.

The question as to whether and to what extent Geller's public demonstrations were genuine or not is of no consequence to these considerations. However, on the basis of a number of investigations it is assumed that among the Geller inspired subjects a reasonable percentage demonstrated PK under conditions which can be accepted with some confidence (Betz, 1978; Wälti, 1978). The numerical success of the Geller inspired tests was partly due to the mass media being able to reach a large proportion of whole populations groups. However large scale tests have been tried before with only limited success (Brier, 1969; Wiklund & Jacobson, 1976).

A more important argument is probably that in 1974 many people believed they saw paranormal events and that there was considerable public acceptance that these paranormal events really occurred. This acceptance was reached under conditions which did not create undue fear of these novel events and which did not demand drastic changes of existing religious or other belief systems. Parise who had demonstrated directly observable PK (Watkins & Watkins, 1974) stated that observing Kulagina, even if only on a film screen, enabled her to continue with PK under conditions which would otherwise have made it impossible for her to do so (Keil et al, 1976). Personal convictions and expectations also play a role in such a situation and could perhaps also account for Parise's experiences.

The assumed PK events reported by the general public, which were at any rate psychologically triggered by Geller, are in many ways in agreement with the theoretical considerations put

forward by Batcheldor (in press) who had been able to produce large scale PK with "sitter groups" of unselected subjects. The main elements which seem to have led to this success were (1) An environment in which normal events cannot be readily distinguished from paranormal ones. (2) Emotional and motivational states within which PK phenomena are regarded as possible and desirable. (3) Personal experiences of apparent but not necessarily real PK and (4) a situation in which the individual participant does not need to feel responsible for any real or assumed PK events if they occur.

Most people who believed that they experienced PK in connection with Geller reported at that time that they thought Geller was fully or partly responsible for the phenomena which occurred in their homes. Most did not believe that they could initiate psi events without help of Geller. (Keil & Hill, 1975).

In 1978, during a period of study leave at the Parapsychological Laboratory, University of Utrecht, the author was initially interested in looking for PK in a laboratory setting while taking account of some of the considerations mentioned above. As a PK target a compass needle was chosen and it was assumed that the sensitivity and small size of such a needle might help to overcome inhibitions and preconceived ideas that it is impossible to produce a PK event. The needle was housed in an apparatus with provisions for cinefilm recordings. An electromagnetic coil in the set-up could be activated by a random event generator such that normal movements of the needle occur in a random time sequence. From a film record it was possible to distinguish clearly between coil activated movements and other movements which may occur.

The Institut für Grenzgebiete in Freiburg kindly offered to make a selection of cases available who had reported PK after the 1974 visits of Geller to Germany and Switzerland and who had responded to subsequent questionnaires (Bender et al, 1975; 1976a; 1976b). Hence it seemed possible to try out the apparatus and certain procedures with a sample population which might succeed with PK more easily than an unselected group of subjects. About 150 cases from 800 questionnaires were selected by the Freiburg Institut and it was initially assumed that his subsample could be restricted to cases who had reported clear distortions of cutlery. It was also expected that quite a number of cases could be found who had experienced such PK events on more than one occasion.

Due to changes in the addresses, illness and other problems the

subsample of 150 decreased to 40. 30 cases were eventually visited by the author. Seven of the excluded 10 cases were not available at the dates on which visits could be arranged. The remaining three were omitted because of travel difficulties.

The author did not consult the questionnaires prior to the first visits in order to avoid any suggestions which might influence the 1978 accounts of the 1974 events. It was interesting to check whether reports given in 1978 agreed with the 1974 questionnaires[1] or whether there were any signs of evaporation (Eisenbud, 1966; West, 1971). As it turned out there were few if any. Most respondents seemed to have vivid memories of the 1974 events. However, since the questionnaires did not have provisions for open ended stories of the events comparison can only be made between the number of respondents who reported distorted cutlery and/or the psi repair of watches in 1974 and in 1978. There was no difference with respect to the watches. Two respondents who only mentioned watches in 1978 had also reported distorted cutlery in 1974. Of these two one had since become more interested in religion and psychic healing and did not regard other psi events as important any more (her own statement in 1978). The second case involved a Swiss who regarded the apparent psi repair of watches as more important than metal bending. An additional third case should perhaps also be mentioned: In the 1974 questionnaire the respondent had reported events in connection with a friend who was then living in the same household but who was not available in 1978. The questionnaire referred to a spoon which bent by only one degree. This spoon was not mentioned in 1978. In 1978 all respondents maintained that they personally experienced or witnessed paranormal events in 1974.

On the basis of this comparison it seems reasonable to suggest that at least in this kind of sample and under the particular circumstances there is little evidence to support the suggestion that a peculiar evaporation of the memory of psi events took place.

The apparatus containing the compass needle was built in such a way that in conjunction with a cine camera it could be used as a portable unit. However, it was impracticable to induce random movements of the needle with the portable set-up and no artifically induced movements were attempted during the visits to the 30 cases.

[1] Eberhard Bauer (Freiburg) provided a summary of the 30 cases from the 1974 questionnaires and his help is gratefully acknowledged.

It was assumed that a compass needle in a sealed housing or container might be expected to be more difficult to move than a needle which can be viewed without a barrier. The apparatus was constructed in such a way that the needle could be directly observed by the subject through a foam padded viewer. The needle could also be photographed and observed through a clear perspex cover. Normal breathing did not induce movements of the needle. Indeed, blowing of air, tapping and other attempts to move the needle artificially were either unsuccessful or appeared to be rather obvious to any observer. However, no claims for perfect control of such factors are made. The strategy was to tighten controls if movements occurred. For instance, the open compass needle in the apparatus could be replaced with one inside a sealed container. No problems were experienced on account of the direct viewing arrangements. With the exception of the first 4 cases an otherwise similar but non-ferromagnetic needle was also available for use with the apparatus.

"With the exception of the first 4 respondents (at that stage the device was not available) a strain gauge was also used. This device[1]) glued with epoxy to a 15 cm strip of 1 x 10 mm stainless steel, will register minute variations in the tension of the steel long before any visible distortion takes place. The gauge consists essentially of a special wire which changes its resistance due to minute compressions or elongations. With the aid of a small amplifier the changes in the resistance can be clearly observed on a meter and filmed if required. With such a sensitive device relatively uniform changes occur because of temperature variations and for other not clearly identifiable but undoubtedly normal reasons. In spite of these normal changes in the system itself it is possible to extract certain indicators of paranormal influence. One such indicator for paranormality is the speed with which these changes occur. A second indicator is the differential rate of change (even if relatively slow) if attempts are made to bend the metal in opposite directions". (Keil, in press.)

For the 1978 study the following procedure was used. The author made arrangements to visit the respondents in their homes. In an opening discussion the 1974 events were reviewed and appropriate notes were compiled after each visit. It was ascertained whether the respondent believed that what he experienced as psi events in 1974, was due to Geller or due to himself (or due to other members of the household). It was further established whether the particular beliefs associated

1) Thanks are due to Dr. Brian Millar, Parapsychology Laboratory Utrecht University, who constructed this device.

with the psi events had changed as well as whether there had been a change in the general attitude towards Geller.

In 1978 all but two of the respondents had come to the conclusion that they themselves or other members of the household had been mainly responsible for the events in 1974. All reported that in 1974 they had regarded Geller at least as partly responsible. Seven of the 1978 respondents spoke for other members of the households who were not available during the author's investigations.

During visits to the two respondents who still maintained that Geller had been the psi source, it was suggested that they themselves must also have some psi abilities because the Geller inspired phenomena seem to work just as well when taped TV programmes are transmitted as during live shows.

It could be argued that a contrary suggestion (that Geller was at least partly responsible) might have created more favourable conditions in 1978 for further psi events. However, since it was quite apparent from the beginning that most seem to have become rather indifferent or even negative in their attitude towards Geller, the author decided to emphasize that the respondents could generate PK on their own.

Unless the attitude to Geller was particularly negative the projection of a cine film of Geller was offered. Only four respondents wanted to see Geller again. The recent adverse publicity of Geller in German and Swiss newspapers probably contributed to this rejection and about one third of the respondents referred to these publications. Some respondents expressed disappointment because they felt that Geller had started to cheat occasionally in order to make as much money as possible. However, in restrospect nobody suggested in any way that the personally experienced paranormal events in 1974 can now be regarded as being due to normal factors.

With one exception (because time was short) all respondents were interested to view a cine film of Kulagina (Keil & Fahler, 1975). Her apparently easy success with compass needles was used to introduce the compass apparatus. With one exception all agreed to try this set-up with the aim of moving the needle. With the exception of the first four cases (when the device was not available) all respondents also agreed to try out the metal bending experiment with the strain gauge.

It was suggested that the movement of a compass needle might be easier than bending a spoon and the author demonstrated prior to the metal bending experiment, that if the metal tended to bend by only a minute amount this was sufficient to be registered by the device. Hence only a fraction of their 1974 success was required. Normal spoons were also offered but most restricted their attempts to the metal with the strain gauge. All observations were carried out under good light conditions.

Without psychological tests and longer and more detailed explorations it is only possible to make a few observations which may be justified when the sample is compared with the general population. Respondents represented a large section of the scale as far as socio-economic status is concerned, perhaps with a somewhat smaller percentage at the lower end of the scale. This can be easily understood if it is kept in mind that the sample had to fill out a questionnaire in 1974 and reply by mail to a circular in 1978. There were about three cases in the sample (and the associated possible psi events were rather weak ones - only one reported metal bending in 1974) who in terms of religious and other belief systems were relatively "far out". It can perhaps be further assumed that the suggested percentage is higher than what one might expect to find in the general population. Nevertheless, the striking impressions was that most of the people were in no way different from what one might expect to find in random samples of the general population. No peculiarities which could in any way be regarded as being common to the sample, could be discovered.

Results with the compass needle: Movements celarly occurred only in connection with one case that was further investigated (Keil and v. Lucadou, in press). On the basis of the 1974 events and on the basis of a PK test in Freiburg there still exists a reasonable possibility that one or more members of the family concerned have psi abilities and that psi events occurred in 1978. However, most and perhaps all of the compass needle movements could eventually be explained on account of the operation of a lift in the same building. As long as this was unknown to all concerned the artifially induced movements of the compass needle had a positive motivational influence on the experimental situation and if any psi events occurred at all in 1978 they would probably not have taken place without this artificial induction.

The assumption that artifically induced movements provided a beneficial framework for the PK tests, is based on a subjective evaluation of the events. Nevertheless, the author tried to make allowances for his own expectations - for some time he did not know himself that the lift was involved - and also tried to employ semi-quantitative behavioural criteria in his evaluation. In particular, it seems highly unlikely that the family would have spent as much time or attempted as many trials (including during periods between visits by the author) without the artifically induced movements.

No clear movements of the compass needle were observed for any of the remaining 29 cases. This suggests that perhaps contrary to expectation a compass needle is quite well suited for experimental investigation. Minute vibrations of the needle could be observed on a number of occasions but even noticable substantial vibrations of the table and of the needle, for instance, on account of heavy traffic, did not produce a turning movement of the compass needle. Unfortunately this is the case when a nonmagnetic needle is used and movements which were observed in 4 cases out of 25 can at best be regarded as only slightly suggestive of psi. Vibrations were not noticeable while the movements of the nonmagnetic needle were observed but this cannot be regarded as an adequate control.

As indicated earlier due to its high sensitivity a strain gauge will register changes which are due to temperature variation and other normal factors. Nevertheless some indication of the relative probability that psi is involved can be gained from the speed with which changes occur and from the directional control which subjects can maintain over a number of trials during which they attempt to bend the metal in opposite directions.

Only two cases were observed who apparently produced rather rapid changes (each within a fraction of a second) in the strain gauge resistance. Changes of this kind were never observed during trial runs with the device. However, neither of the two respondents was able to repeat the fast changes which in one case happened three times and in the other only once. Consequently it is difficult to rule out the possibility that the strain gauge was activitated by unknown normal factors.

A number of subjects seem to have had success in obtaining strain gauge changes which were in agreement with the desired bending directions. However these trials were carried out while

in most cases a relatively slow and uniform one directional change (presumably due to temperature change) took place and there was some difficulty and some unavoidable subjectivity in assessing the relative success or otherwise of various subjects. Four cases involving a total of 17 trials of attempts in opposite directions were rated as strong. Four further cases involving 9 trials were rated as suggestive but weak. If it is assumed that some PK occurred among these 8 cases then the higher success rate (compared to the compass needle) could be attributed to the higher rate of change on account of normal variables of which probably some were experienced by some subjects as paranormal. On the other hand it cannot be ruled out with certainty that most of the apparent psi success was due to fortuitous combinations of a number of normal variables and due to faulty observations. However, observations of the device with unsuccessful subjects provided some confidence that successful trials were not simply due to misperceptions or other simple normal factors.

All respondents were asked whether they had other psychic experiences or abilities with a view to investigate any claims which can be reasonably tested. While a number of spontaneous ESP experiences and one case involving paranormal bending did not seem to be suitable for reasonable test procedures, one claim of water divining was further investigated.

The son A.H. aged 14 in 1978, seems to have been mainly responsible for possible PK events in 1974. The father was not interested in parapsychology but took water divining and metal divining seriously because it was used in a practical way in the building industry in which he was working. In that area of Germany apparently many building contractors who excavate trenches use divining rods to locate underground cables and water pipes. A.H. had tried his hand at locating metal and water and claimed to be quite sure of his ability to locate water or metal underground. He admitted to some difficulties in distinguishing between the two. The divining rod used by A.H. and apparently by building contractors in this area is somewhat different from those usually referred to in the literature. The rod consists of two separate halves, one held in each hand. The two halves are similar and usually made of 1.5 to 3 mm. metal wire. A.H. used copper. Each half is simply made by bending a right angle into a length of 100 cm. wire such that one side is 20 cm. and the other 80 cm. long. The 20 cm. side is held vertically and the 80 cm. side horizontally. Initially both rods are held in such a way that

the two 80 cm. sides are parallel and point ahead of the diviner. As the diviner moves slowly over an area the two long sides turn towards each other as metal or water is approached.

This movement can be explained by normal means if it is assumed that the diviner - probably without awareness - moves the short ends of the wire into a slightly non-vertical position. The question remains whether a subject is able to locate metal or water more often than would be expected by chance. This is difficult to estimate in an uncontrolled situation. Building contractors may react to normal cues without being aware of it and even if someone is apparently highly successful no psi may be involved. Nevertheless, divining rods also provide a situation in which possible paranormal events are readily accepted partly because they are not necessarily experienced as paranormal at all and partly because the responsiblity is removed to the divining rod which is supposed to move because of unknown radiation or suchlike from the metal or water. When A.H. agreed to try out tests which could be evaluated to some extent the author encouraged A.H. to go ahead and later returned for a further visit and some joint tests.

In the author's absence A.H. tried the following procedure based on the author's suggestions. A heavy piece of metal could be hidden underneath a ladder in three different positions. The ladder was placed on two low stools and this arrangement was covered to the ground with a large blanket. The three target areas were marked out with chalk. While A.H. waited inside a shed which did not enable him to view between trial preparations accidentally, another member of the family moved the metal under one of the three positions. The target positions were "randomly" selected but the procedure is not quite clear. The author had suggested that dice should be thrown. In a sequence of 20 trials A.H. had 16 hits.

During the author's second visit the test was repeated. Some care was taken that no normal cues were provided through sounds or through changes in the fall of the blanket. (After the metal has been placed in the particular target position the blanket was each time pulled straight from the same position). Each trial was announced in a similar way with a metal gong (quite separate from the target metal) and the target sequence was determined with a miniature portable roulette wheel.

His score was actually the same as the one which he achieved in the author's absence (16 hits in 20 trials). However during 30 further trials only chance results occurred and when A.H. "checked his abilities" by using the rod when the metal was placed with his knowledge in one position, he reacted nevertheless to the empty positions as well. Claiming that water interfered, he wanted to continue the tests some distance from his home where there was apparently less interference but because of time and travel limitations that was not possible. The case cannot be regarded as a strong one for providing evidence for psi. The controls during the test situation were by no means as strict as might be desirable particularly if one assumes that all psi subjects are amateur or professional magicians. No indications of any kind occurred that such suspicions were justified and it does not seem unreasonable to claim that the test results were suggestive of psi. The procedure worked well and with modifications, depending on particular circumstances, it should probably be kept in mind for future use. If the possibility that in such a situation a subject may react to magnetic field changes is taken seriously, an artificial magnetic field would have to be introduced as well.

The 30 visits to localities all over Germany and Switzerland were time consuming since most respondents were working adults with various commitments. Often long journeys could not be avoided even when a number of cases were located in one area because suitable successive appointments could not be arranged. Nvertheless sufficient time was left after the field investigations to test a sample of comparable size under laboratory conditions. Unfortunately on account of university vacations and other local circumstances only 5 subjects were available for laboratory investigations. This number is too small for a meaningful comparison and these experiments can only be regarded as a pilot study for future laboratory work which hopefully will be continued in Tasmania. No portable random event generator was readily available at that time and the small number of subjects did not warrant more complicated arrangements. Nevertheless, artificial movements of the compass needle were introduced on the basis of a random number table. No movements of the compass needle were observed which could be labelled as psi events. At the time of the experiments no test rooms were available which could be regarded as providing a psychological setting likely to encourage psi. In spite of the artificial movements it was difficult to create an environment in which subjects were likely to experience such movements as real psi events. Under the circumstances the lack of results with the compass needle is not altogether

surprising. However, the main problem was the small number of subjects and consequently it is not really possible to evaluate the experiments which were carried out in a laboratory setting. Movements of the nonmagnetic needle occurred in two cases but vibrations on account of heavy traffic were noticeable in the test room at that time. The strain gauge experiments produced only weak suggestions of psi events.

In retrospect two factors made the field investigations less advantageous than originally expected. The first was that at least for the time being the original plan to carry out laboratory experiments, could not be followed to a meaningful extent. The second was that the selected sample of 30 cases showed less psi potential than was expected when this matter was first discussed in Freiburg. Less than half of the 30 respondents reported metal bending in 1974 and if it is taken into account that some cases involved other members of the household, only one third had actually experienced metal bending in 1974. Unfortunately the 30 cases did not include any claims for PK experiences on more than one occasion.

There was no way of clearly distinguishing between possible psi metal bending events and those which were reported because of normal intentional or unintentional activities or misperceptions. While some of the events associated with watches may have had a psi component the probability that his was the case is presumably much lower because many discarded watches may start to work again (at least for a while) simply because they are moved or warmed up. During the 1978 investigations with the strain gauge there was with one exception no indication that respondents attempted to produce distortions by normal means. The one exception refers to a subject who participated with great enthusiasm and who without being aware of it, held the metal in such a way and applied sufficient normal pressure that the resulting distortions can safely be regarded as being due to normal factors. Hence her 1974 report about distorted cutlery must be regarded as highly doubtful with respect to psi.

On the positive side some details which emerged in 1978 suggest that in some cases at least unconscious manipulations were unlikely. In one case two links of a watch chain had distorted and fused together in a way which seemed impossible to achieve except with special tools and a high temperature flame. Apparently a watchmaker who inspected the chain admitted that he would have the greatest of difficulties in producing a similar result. In another case

two spoons had bent inside a closed dishwasher by at least 90° (Keil and v. Lucadou, in press). No malfunction of the machine was ever noticed and it is unlikely that the machine can be regarded as a reasonable explanation for the reported distortions. In a third case a fork which could still be inspected in 1978, had bent sideways and was distorted along the lines of a small circular shape in such a way that it would be difficult to achieve this even with the help of tools.

If any of the events triggered by Geller are taken seriously at all it is probably not unreasonable to assume that the sample of 30 contained about 5 cases who probably produced psi phenomena in 1974. The rest of the sample still differed from the general population as most believed they had experienced or witnessed psi in 1974. With some variations most or all also had a higher expectation that psi events might occur again.

The experience of visiting the 30 cases was a fascinating and positive experience for the author. Some reports of spontaneous events known to some of the respondents included an impressive and detailed account of a Poltergeist case. It occurred many years ago in the same village in which the respondent was then living. In another case some personal events during the last world war has been critically and reasonably evaluated and still suggested a psi component.

One was reminded that if psi can be extracted in the laboratory it must surely exist in real life situations and once certain inhibitions are removed it is possible to find a good deal of evidence which suggests that psi is experienced in real life situations. More important from an experimental point of view is the fact that one of the 30 cases succeeded to move the compass needle through intentional or unintentional normal manipulations. There is no doubt that it is difficult to detect the methods with which an experienced magician can pretend to produce paranormal events (Johnson, 1976). On the other hand preoccupation with trickery of public performers and the widescale rejection of real psi events by magicians can create a climate within which parapsychologists become inclined to expect trickery whenever psi is mentioned. Yet compass needles which any 5th rate magician would regard as easy to manipulate were not subjected to trickery by any of the more than 30 people who had reported psi experiences before and who from the point of view of the critic must have fooled themselves through unconscious manipulations or others through trickery. Yet apart from one person discussed previously

who probably bent metal by normal means without being aware of it, there was not even a sign of unintentional manipulation throughout the whole investigation.

From the experience gained the author certainly feels encouraged to continue with similar field investigations and to recommend them to others. With the ever increasing miniaturization of electronic equipment the degree of control which can be achieved in the field may actually be higher than in the laboratory in which the psychological setting required for a particular person cannot be adequately controlled. This is not to say that field experimentations are preferable but rather that they ought not to be neglected if opportunities arise.

REFERENCES

Batcheldor, K.J. PK in sitter groups. Psychoenergetic Systems, in press.

Bender, H., Hampel, R., Kury, H., Wendlandt, S. Der "Geller-Effekt" I Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1975, 17, No. 4, 219-240.

Bender, H., Hampel, R., Kury, H., Wendlandt, S. Der "Geller-Effekt" II Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1976, 18, No. 1, 1-20.

Bender, H., Hampel, R., Kury, H., Wendlandt, S. Der "Geller-Effekt" III Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1976, 18, No. 2, 105-116.

Betz, H.D. Unerklärte Beeinflussung eines elektrischen Schaltkreises durch Silvio. Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1978, 20, No. 1, 47-50.

Brier, R.M. A mass school test of precognition. J. Parapsychol. 1969, 33, 125-135.

Eisenbud, J. The problem of resistance to psi. In Roll, W.G. (Ed). Proceedings of the Parapsychology Association. 1966, No. 3, 63-79.

Johnson, M. Some reflections after the P.A. Convention. E.J.P. 1976, 1, No. 3, 1-5.

Keil, H.H.J., Fahler, J. Nina S. Kulagina: A strong case for PK involving directly observable movements of objects. E.J.P. 1976, 1, No. 2, 36-44.

Keil, H.H.J. A follow-up study of 30 possible PK cases. R.I.P. 1978, in press.

Keil, H.H.J., Herbert, B., Ullman, M., Pratt, J.G. Directly observable voluntary PK effects: A survey and tentative interpretation of available findings from Nina Kulagina and other known related cases of recent date. Proceedings SPR, 1976, 56, 197-235.

Keil, H.H.J. & Hill, S. Mini Geller PK cases. R.I.P. 1974, 1975, 69-71.

Keil, H.H.J., von Lucadou, W. "Psychokinese oder Pseudopsychokinese: Eine Feldstudie. Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, in press.

Millar, B. The observational theories: a primer. E.J.P. 1978, 2, No. 3, 304-332.

Wälti, B. Die Silvio-Protokolle 1976-1977. Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1978, 20, No. 1, 1-46.

Watkins, G.K., & Watkins, A. Apparent psychokinesis on static objects by a "gifted" subject: A Laboratory demonstration. R.I.P., 1973, 1974, 132-134.

West, D.J. Reasons for the continuing doubt about the existence of psychic phenoma. In Angoff, A. and Shapin, B. (Eds). A century of psychical research: The continuing doubts and affirmations. N.Y. Parapsychology Foundation, 1971, 155-159.

Wiklund, N. & Jacobson, O. A public experiment with precognition. Journal S.P.R., 1976, 48, No. 769, 293-300.

H.H.J. Keil
University of Tasmania
Box 252 C, GPO Hobart, Tasmania
Australia 7001

CONFORMANCE BEHAVIOR INVOLVING ANIMAL AND HUMAN SUBJECTS

Elmar R. Gruber

Institut für Grenzgebiete der Psychologie und Psychohygiene

INTRODUCTION

The experiments reported in this paper are based on Stanford's concept of conformance behavior (Stanford, 1977a; 1977b; 1978). The purpose of this study was to show if conformance behavior involving living systems does in fact occur. The hypothesis according to experiments done in this direction is, that living systems will be quite susceptible to conformance behavior because of their complexity and plasticity (Braud, 1978). The experiments reported here were designed partly as experiments of "biologically-generated" pre-recorded targets (cf. Schmidt, 1976) and partly as a replication and extention of Braud's (1978) findings. Successful replications of Braud's findings have already been reported elsewhere (Gruber, 1979). The last part of the experiments consisted of studies of conformance behavior involving human subjects as the relevant biological system for the generation of "random events".

Let us briefly sum up Stanford's concept: For conformance behavior to occur we need: (i) a disposed system (usually an organism with a need or other disposition); (ii) a random event generator (that is a source of incompletely determinate alternative states or events); (iii) a contingent relationship such that the probability of a favourable event (according to the disposition) is linked to the REG outputs. Conformance behavior is said to occur when these elements are present and the REG outputs become biased so as to increase the probability of the favourable event.

GENERAL PROTOCOL

12 experiments concerning conformance behavior in rodents and humans were performed. Animals involved in the study were four gerbils (Meriones unguiculatus), each participating in several experimental sessions. 30 volunteer subjects participated in experiments 9 through 12 as well as an additional subject, who served as co-experimenter during

some parts of the remaining experiments. 20 additional volunteer subjects participated in experiments 1 through 8, each contributing two experimental sessions. Of the 50 subjects (excluding the experimenter and the co-experimenter) 32 were females and 18 males. Experimental hypotheses did not include sex differences. The only hypothesis involved in both animal and human studies was, that conformance behavior would show up between locomotor activity of animals and human subjects according to a specified intention of a motivated distant observer.

The experimental set-up for the animal studies was as follows: The cage with the animal was situated in a small room. The experimenter monitoring the tape-equipment was situated in another room, separated by the room with the animal by yet another small room. The subject given feedback of the animal's activity was placed in another room, adjoining the room with the experimenter.

Each session involved twenty 30-second epochs of which ten were control periods and ten conformance periods. During the control periods the feedback mechanism was turned off and the subject simply relaxed without intentions for the biological system. During the conformance periods, the subject tried to maintain a specific intention for the biological system to behave in a specified manner. Equal numbers of conformance and control epochs were generated prior to each session by a randomization procedure employed in similar studies (Braud, 1978; Gruber, 1979a). Unlike the Braud (1978) and Gruber (1979a) studies, in which single mean $\underline{t}$ tests were used for statistical comparison of experimental and control periods, Mann-Whitney $\underline{U}$ tests were calculated between the 10 conformance and 10 control periods for each subject or session. Similar to a recent study (Gruber, 1979b) the probability values thus derived were combined across subjects by the chi-square method recommended by Guilford (1956). Because of positive results in former experiments of this kind (Braud, 1978; Gruber 1979a; 1979b) one-tailed tests were chosen for the animal studies, while two-tailed tests were employed in the study with humans. The 5% level was chosen as significance level.

Precautions against conventional sensorimotor and conventional energetic and informational interaction between the biological system, the experimenter(s) and the subject were taken.

The different kinds of biological systems used were the following:
(i) locomotor activity of a gerbil (experiments 1-4)
(ii) pre-recorded locomotor activity of a gerbil (experiments 5-8)
(iii) locomotor activity of a human subject (experiments 9-10)
(iv) pre-recorded locomotor activity of a human subject (experiments 11-12).

SPECIAL PROCEDURES AND RESULTS

Experiment 1 and 2

Both experiments one and two were designed as replication studies of the experiments of William Braud (1978). The experimental procedure was as follows: the subject (observer) was brought in, talked to, shown the experimental animal that was put into the experimental cage shortly before the subjects arrival and finally the purpose of the study was described to him. All subjects were known to the investigator and had different attitudes about psi in general and the present study in particular. Nevertheless, all subjects expressed their motivation and interest in the experiment. Afterwards the subject was brought into a room where he was seated in a comfortable reclining chair. A short paper with instructions about the experiment was handed to the subject (see Appendix). Then the subject put on the headphones and the experimenter left the room and moved into the adjoining room, where the tape-recording equipment was located. The locomotor activity of gerbils was mesured by allowing the animals access to an activity wheel after five minutes spent in the holding cage attached to the wheel. The gerbil was deprived of the activity wheel for at least 24 hours before each experiment. Each revolution of the activity wheel activated an automatic counting device while at the same time producing a clearly audible click. The clicks were recorded on magnetic tape while they were simultaneously fed to the observer in the other room via headphones .

Immediately before each session, the experimenter shuffled a pack of twenty cards containing ten black and ten white cards, twenty times and placed the shuffled pack upside down in front of him. By taking a card from the pack before each 30-second epoch, it was determined whether the following epoch was to be a conformance or a control period (a white card indicated a conformance period and a black card a control period). The tape-recorder on which the clicks

produced by the wheel's revolution were recorded was shielded from the experimenter by a box so that he could not observe the display of the VU-meters giving visual feedback for each recorded click. Identical to previous experiments (Gruber, 1979a; 1979b) he was provided with a device with three buttons, one for the turning on of the tape, one for turning the headphones of the observer on and off and one for the production of a one-second buzzer sound also fed to the subjects headphones and indicating the start of a conformance epoch. When the second button was pressed the observer could hear the clicks produced by the revolutions of the activity wheel at a well-audible volume, pre-stabilized before the beginning of the experiment. The experimenter was furthermore provided with a digital watch which started working at the onset of the tape, and with another pair of headphones which served as an additional auditory shielding device and through which only the buzzer sound could be heard.

The subject was informed by means of the instruction paper that after every thirty second epoch there would be a fifteen second rest period in which he would be asked simply to relax and not think about the gerbil. The same applied to control periods. Each conformance period was indicated by a one-second buzzer sound two seconds before the end of the intervening rest period. During the experimental (conformance) epochs, the observer was asked to maintain an intention to enhance the click rate which he could hear through the headphones, while visualizing the experimental animal running quickly in the wheel. If a control period was to follow the rest period, no buzzer sound was provided and the volume was turned off (during rest as well as control periods).

Before the onset of tape-recording and during the rest periods, the experimenter took the following card from the pack to ascertain whether a buzzer sound should be provided, thus indicating a conformance epoch (determined by a white card) or if no such sound should be provided and a control period should follow (determined by a black card). At the end of the experimental session the experimenter pressed the buzzer three times. After the experiment, the subject left without contact with either the experimenter or the co-experimenter. This was also indicated in the instruction paper in order to avoid sensory cueing about possible beliefs on the part of the observers about their

alleged success or failure, which could lead to some sort of secondary feedback on the experimenter and psi experimenter influence on the outcome of the experiment.

Both experimenter and co-experimenter were trained extensively in handling the experiment and in particular about the use of the cards as a randomization procedure and the indicating of the epochs with accompaining turning off or on of the volume in the observers headphones. A person not aware of the purpose of the experiments checked and counted the twenty 30-second epochs of clicks and performed the statistical analysis.

For experiment 1 (investigator as observer), the chi-square of combined probabilities of the U-values of each session was $\chi^2 = 17.42$, df=2k=20, p = .70, one-tailed. For experiment 2 (ten unselected volunteers as observers, each contributing one session) the chi-square of combined probabilities was $\chi^2 = 34.91$, df=2k=20, p = . 025, one-tailed. The chi-square of combined probabilities across all 20 sessions of experiments 1 and 2 was $\chi^2 = 52.33$, df=2k=40, p = .10, one-tailed.

Although the results for the unselected volunteers is significant in the predicted direction (experiment 2), the chi-square of combined probabilities for both experiments does not reach the significance level.

Experiment 3 and 4

Experiments 3 and 4 were identical to experiments 1 and 2, except that instead of the activation wheel, another measure of locomotor activity of the gerbils was used. The gerbil was deprived for at least 24 hours before each experimental session of intense locomotor activity by being placed in an extremely narrow cage. Shortly before the experiment started, the gerbil was placed in a larger cage which was divided into two identical halves by a light-beam (cf. Gruber, 1979a). Each crossing of the light beam by the animal resulted in a click produced by the photo-electric device when entering and when leaving the light-beam. Identical to experiments one and two, the click sequence was presented as auditory feedback to the subjects. Since the enhancement of the click-rate in this experimental setting using mice has been demonstrated successfully elsewhere (Gruber, 1979a), some sort of conceptual replication (Sargent, 1979) was tried, asking

the observers to maintain the intention to reduce the click-rate and visualize the gerbil staying still in some part of the cage. Since conceptual replications of this kind, although with a different experimental set-up, were performed successfully (Gruber, 1979a) it was decided to use one-tailed tests here as well.

For experiment 3 (Investigator as observer), the chi-square of combined probabilities of the U-values of each session was $\chi^2 = 33.82$, df=2k=20, p = .05, one-tailed. For experiment 4 (ten unselected volunteers each contributing one session) the chi-square of combined probabilities was $\chi^2 = 31.52$, df=2k=20, p = .05, one-tailed. The chi-square of combined probabilities across all 20 sessions of experiments 3 and 4 was $\chi^2 = 65.34$, df=2k=40, p = .01, one-tailed.

Thus, the locomotor activity of gerbils in the described direction conformed significantly with the intentions of motivated distant observers.

Experiments 5 and 6

The experimental set-up of experiments 5 and 6 was identical to experiments 1 and 2. However, the twenty 30-second epochs, as well as the intervening rest periods together with the buzzer sounds, were recorded on tape previous to the presentation to observers. No person was able to hear the click sequence until between one and three days after the recording, when, as usual, it was presented through headphones to observers. The co-experimenter performed the ten sessions of experiment 5 and the ten unselected volunteers who participated in experiment 3 performed each one session. Neither the co-experimenter nor the subjects knew that these targets were pre-recorded. Since experiments of this kind with pre-recorded target material yielded positive results in previous experiments (Gruber, 1979a; 1979b) once more one-tailed tests were used.

For experiment 5 (co-experimenter as observer), the chi-square of combined probabilities for each session was $\chi^2 = 37.02$, df=2k=20, p = .0125, one-tailed. For experiment 6 (ten unselected volunteers each contributing one session) the chi-square of combined probabilities was $\chi^2 = 23.91$, df=2k=20, P = .30, one-tailed. The chi-square of combined probabilities across all 20 sessions with pre-recorded biologically generated targets was $\chi^2 = 60.93$, df=2k=40, p = .025, one-tailed.

Thus a conformance effect could also be found in a time displacement study.

Experiments 7 and 8

Experiments 7 and 8 were identical to experiments 3 and 4 as far as the experimental set-up was concerned. The procedure was identical to experiments 5 and 6, thus presenting pre-recorded targets and the intention of the observers to reduce the click-rate and hence the locomotor activity of the gerbil.

For experiment 7 (co-experimenter as observer), the chi-square of combined probabilities for each session was χ^2 = 20.25, df=2k=20, p = .50, one-tailed. For experiment 8 (ten unselected volunteers as observers) the chi-square of combined probabilities was χ^2= 28.50, df=2k=20, p = .10, one-tailed. The chi-square of combined probabilities of all 20 sessions of pre-recorded biologically generated targets was χ^2 = 48.75, df=2k=40, p = .30, one-tailed.

Thus a conformance effect for the reduction of the click-rate of pre-recorded targets could not be found.

TABLE 1

Results of gerbil experiments

Experiment	chi-square	df	p-value
1	17.42	20	.70
2	34.91	20	.025
3	33.82	20	.05
4	31.52	20	.05
5	37.02	20	.0125
6	23.91	20	.30
7	20.25	20	.50
8	28.50	20	.10

all p-values one-tailed
significance level: .05

Experiments 9 and 10

In experiments 9 through 12 the biological system involved for the generation of "random events" was the locomotor activity of a human subject. For these experiments a special room was constructed consisting of four wooden walls each 2.50 meters long and 2.00 meters high. The square-room was divided into two identical halves by a light-beam at the height of 30 centimeters from the ground. Subjects were shown the experimental room while the photo-electric device was shielded. The eyes of the subjects were covered by opaque eyeglasses and they were provided with headphones through which pink noise was fed during the whole length of the experiment. Subjects were psychology students and they were told that this was a psychological experiment about special performance under extreme conditions or states. To make them believe in this, they had to complete a questionnaire about their feelings and experiences during the "random walk" in the experimental room. Before they entered the room they were asked to walk around in an as "random" as possible fashion and to avoid any patterning effects or habituations as to the direction of walking as well as to the speed and the rhythm. The subjects were allowed a three-minute training phase before the actual experiment started, which was unknown to them. During this period, they usually became acquainted in a very short time with the properties of the room which was covered on the inside with sponge rubber since the subjects often hit the walls. In addition to the shielding of the eyes, the experimental room was usually left in complete darkeness.

Meanwhile, an observer was provided with a pair of headphones in a room separated from the experimental room by four adjoining rooms. He was seated in a comfortable reclining chair and the purpose of the experiment was explained to him. Identical to the previous experiments with the gerbils, the observers (who did not participate in experiments 1 through 8 and were almost all unknown to the experimenter) were asked to maintain the intention to enhance the clickrate fed through the headphones and to visualize the subject in the experimental room moving frequently across the light-beam. The click-rate was produced by the passing of the subject's legs across the light-beam with a corresponding click when entering the light-beam and one when leaving it. Hence the rate and intervals of the clicks was determined by a very complex procedure which should be favourable for conforming with a more structured and less random system related to it.

For both experiments 9 and 11 with the experimenter and the co-experimenter respectively as observers, two-tailed tests were employed, as this was the first time experiments of this kind were performed and effects in both directions were to be expected.

The experimenter was situated in the room with the shielded tape-recorder and monitored the experiment indicating conformance and control epochs of 30-seconds each by the buzzer sound (as in experiments 1 through 8, ten epochs each). After the subject in the experimental room had started his three-minute "random walk", the real experiment started with the observer's intention to produce conformance behavior between the "random walk" of a human and his specific disposition. 20 persons participated in this experiment, ten being subjects in the experimental room and ten being observers. The same subjects participated in experiments 11 and 12 as well. During each experimental session, there was only one subject in the experimental room, just as in experiments 1 through 8 there was only one gerbil in the cage.

After the experimental session was over, the observer left without further contact with the experimenter, who then went to the experimental room and took the headphones and the eyeglasses from the subject. He then handed the subject a questionnaire about the feelings and experiences he had during the walk in the room. The subjects were also asked about their estimates of the distribution of the percentage of their stays in unspecified parts of the room. The intermediate relay producing the clicks was situated in an accoustically well-shielded box.
For experiment 9 (investigator as observer) the chi-square of combined probabilities for each session was $\chi^2 = 40.25$, df=2k=20, p = .01, two-tailed. For experiment 10 (ten volunteers as observers each contributing one session) the chi-square of combined probabilities was $\chi^2 = 32.50$, df=2k=20, p = .05, one-tailed. The chi-square of combined probabilities across all 20 sessions of experiments 9 and 10 was $\chi^2 = 72.75$, df=2k=40, p = .00125, one-tailed.

Thus a significant conformance effect was found between the intentions of motivated distant observers and the locomotor activity of human subjects.

Experiments 11 and 12

Experiments 11 and 12 were identical to experiments 9 and

10 but the click-sequences presented to the observers were pre-recorded between one and six days prior to the playing back. Observers, who were the same as in experiment 10, did not know that the click-rate had been pre-recorded and thought that the experiment would be the same as in experiment 10. Subjects for the "random walk" all participated for the first time in experiment 12. No subject for the "random walk" participated in two experimental sessions.

For experiment 11 (co-experimenter as observer, the chi-square of combined probabilities for each session was χ^2 = 31.30, df=2k=20, p = .10, two-tailed. For experiment 12 (ten volunteers each contributing one session) the chi-square of combined probabilities was χ^2 = 20.24, df=2k=20, p = .50, one-tailed. The chi-square of combined probabilities across all 20 sessions with pre-recorded biologically generated targets was χ^2=51.54, df=2k=40, p = .10, one-tailed.

Thus a conformance effect concerning the "random walk" of human subjects and the intentions of motivated and disposed distant observers in a time-displacement setting could not be found.

TABLE 2

Results of human experiments

Experiment	chi-square	df	p-value
9	40.52	20	.01+
10	32.50	20	.05
11	31.30	20	.10+
12	20.24	20	.50

+ two-tailed

significance level: .05

DISCUSSION

In the experiments reported here conformance effects have been found in some cases, also in time-displacement settings, for conformance between gerbils and human observers and between humans and human observers.

It may have been possible that the animals could hear the clicking produced by their own behavior on the intermediate relay even though it was accoustically well shielded. The results suggest that this might not have been of damaging influence, although the exact role of hearing the clicks by the animals has yet to be analyzed (in the case of the activation wheel study, the shielding was not as good as in the light-beam setting).

In the case of the conformance of a human "random walk" with intentions of an observer, the intermediate relay producing the clicks was accoustically well shielded and moreover, the persons were fed pink noise through headphones. In this study some interesting features were found post hoc suggesting more such experiments in the same direction. Firstly, I did not expect a significant conformance effect in this study: primarily for the reason that it indeed seems that the subjects are not able to produce really "random" walks, a fact which, according to observational theories should produce some difficulties for, say, a PK influence. The degree of randomness of the walk was not controlled in this experiment, on the one hand because it was only intended to serve as a pilot study to show if conformance behavior in such a setting occurs at all, and on the other hand because observations other than those by the experimental observer, were avoided. In spite of this, some sort of control of randomness in relation to the degree of conformance should be provided in further studies.

Another interesting fact that was found in the human studies was that some of the subjects walking in the square room quite rapidly fell into an altered state of consciousness. This was clear from the answers they gave in the questionnaire, which was initially only provided to make them believe in a psychological experiment. The subjects who were deprived of visual input and provided with a homogenous accoustic stimulation together with motoric activity reported in some cases, distortions in time and space sensing. One subject even fell repeatedly to the ground as a result of complete loss of spacial relation-sensing. Others reported of vivid hypnagogic hallucinations. One girl seems to have had very deep emotional experiences together with the images of falling into a deep dark chasm. When the experiment was over she was totally disoriented and refused to re-enter the experimental room which she termed the "hall of demons". Others found the walk under

these strange conditions pleasant, even if some thought that they had to walk for too long. This might also be a psychological factor, since most of the subjects felt that they were in the room much longer than they actually were.

During the time-displacement study, which was not significant for the whole experiments 11 and 12, there were six subjects (five of which were female) who reported very distinct features of having been in an altered state of consciousness during the experiment. A post hoc analysis of these six subjects revealed that the chi-square of combined probabilities of their respective U-values was significant (χ^2 = 27.30, df=2k= 12, p= .01, one-tailed). Although this finding was post hoc it is an interesting result, suggesting that conformance should be better when an ASC is observed than when the subject is normally aware of his acts and his special environment. The loss of orientation during the ASC may be constitutive either for being less aware of one's own behavior (thus favouring concepts of the observational theories in this respect) or for helping in the production of "more random" activity (thus leading to a better conformance effect according to the underlying concept).

It is planned in future studies to explore this field more extensively and to administer special questionnaires asking about ASC's in particular, together with measures of randomness in order to test these conjectures.

Non-parametric statistical analysis was performed for all the experiments reported here as in another study (Gruber, 1979b), since parametrical statistics, as employed in yet other studies (Braud, 1978; Gruber, 1979a) do not seem to be appropriate for this kind of experiment.

One interesting problem in interpreting the obtained results lies in the fact that both the experimenter (in these and previous experiments), as well as the co-experimenter, (in one experiment) obtained significant results when acting as observers. Since in the framework of Stanford's model no feedback is necessary (as it is for observational theories) a conformance effect between the experimenter's intentions and the biological REG might be possible. Moreover, an experimenter who designs the study might be (and in most cases is) the most motivated person participating in the experiment and hence the most likely person responsible for conformance effects. Although the results were not as good

as in a previous experiment (Gruber, 1979a) there are still a number of significant results which cannot ultimately be assigned to a subject or an experimenter. The problem of coping with the experimenter effect still rests as one of the most crucial issues (Millar, 1979; Spinelli, 1979; Stanford, 1979).

However, it seems from these experiments that research in this direction might be fruitful in the future, especially concerning study with humans.

ACKNOWLEDGEMENTS

Thanks are due to Brian Millar and Walter von Lucadou for extensive discussions and helpful suggestions during the preparation of this study. Thanks are also due to all the friends who helped in conducting the experiments and especially to the Parapsychology Foundation, Inc., New York, who made this study possible by way of a generous grant.

REFERENCES

Braud, W., Conformance behavior involving living systems. Paper presented at the 21st P.A. convention, St. Louis, 1978.

Gruber, E.R., A study of conformance behaviour involving rats and mice. Paper presented at the 3rd SPR conference, Edinburgh, 1979(a).

Gruber, E.R., Conforming of pre-recorded group behavior with disposed observers. Paper submitted to the 22nd P.A. convention, Berkeley, 1979 (b).

Guilford, J., Fundamental statistics in psychology and education. New York: McGraw-Hill, 1956.

Millar, B., Psi: experimenter effect or experimenter fraud? Paper presented at the 3rd SPR conference, Edinburgh, 1979.

Schmidt, H., PK effect on pre-recorded targets. Journal of the American Society for Psychical Research, 70, 1976, 267-291.

Spinelli, E., The strange case of the experimenter effect. Paper presented at the 3rd SPR conference, Edinburgh, 1979.

Stanford, R.G., Experimental psychokinesis: A review from diverse perspectives. In: Wolman, B.B. (Ed.), Handbook of parapsychology. New York: Van Nostrand Reinhold, 1977, 324-381 (a).

Stanford, R.G., Conceptual frameworks of contemporary psi research. In: Wolman, B.B. (Ed.), Handbook of parapsychology. New York: Van Nostrand Reinhold, 1977, 823-858 (b).

Stanford, R.G., Toward reinterpreting psi events. Journal of the American Society for Psychical Research, 72, 1978, 197-214.

Stanford, R.G., Are we shamans or scientists? Paper submitted to the 22nd P.A. convention, Berkeley, 1979.

Elmar R. Gruber
Institut für Grenzgebiete der Psychologie und Psychohygiene
Eichhalde 12
D 7800 Freiburg i.Br.
W-Germany

APPENDIX

Instruction paper for the activitywheel setting experiments with gerbils.

You are participating in a parapsychological experiment, which, with a very similar procedure yielded significant results. As was shown to you by the experimenter, you should try hard to enhance the locomotor activity of the gerbil running in the activity wheel, by trying to enhance the clickrate which you will hear through the headphones and which corresponds to the animal's activity. There will be 20 thirty-second periods in which you will hear the clicks (experimental periods) and or hear no clicks (control periods). After each period there will be a 15 second rest period. During the experimental periods, which will be indicated by a buzzer sound two seconds before they commence, which you will also hear through the headphones, you should try to imagine the gerbil running quickly in the activity wheel. You should have no such intentions to enhance the locomotor activity of the animal when no buzzer sound is provided, thus indicating a control period or a rest period. During these periods you should simply relax and not think about the animal. You will hear no clicks at all during these periods.

There will be three consecutive buzzer sounds indicating

the end of the experimental session. Please leave the experimental room after the session without further contact with the experimenter, who will contact you later. Thank you for participating in this experiment.

A PARTLY INDEPENDENT REPLICATION OF INVESTIGATIONS OF CASES SUGGESTIVE OF REINCARNATION

Satwant Pasricha Ian Stevenson

University of Virginia

Children who claim that they lived before and can remember details of the claimed previous life can be found in many countries of the world, including those of West Europa and North America. They are particularly common, or more commonly reported, in India compared with other countries. Indian children who claim to remember previous lives usually state verifiable names (of places and persons) and details of events associated with the person whose life is being remembered.

One of us (I.S.) has investigated cases of this type in India and elsewhere since 1961. He has published approximately fifty reports of such cases, most of them in detail (Stevenson, 1966/1974, 1975, 1977a) but a few in summary form (Stevenson, 1960, 1977b). Data on a total of approximately 1,700 cases have been collected and partly analyzed. Among the cases of this collection, more than 200 have occurred in India (Pasricha and Stevenson, 1977; Stevenson, 1966/1974, 1975; Stevenson et al., 1974).

The persons immediately concerned in a case of the reincarnation type - the subject's parents, siblings, and other members of his family and the family of the deceased person (previous personality) whose life the subject claims to remember - usually adopt rather simplistic attitudes to the case in their midst. This is not to say that they accept it uncritically. Indeed, they are likely to interrogate the subject intently and even to mislead him in an effort to expose some ignorance or possible deception on his part. In general, however, they consider only two possible interpretations for a case: a hoax and reincarnation.

For more detached observers, several other interpretations deserve at least as much attention as the two just mentioned, if not more. We should exclude all normal explanations before we consider paranormal ones seriously, and we should accept

reincarnation as the best interpretation only after we have excluded all other possible explanations. For each of these we have some independent evidence, whereas for reincarnation we have only the evidence presented by these cases.

In our opinion, one of the most important normal explanations for cases suggestive of reincarnation is source amnesia (cryptomnesia). This may occur in some cases, but it cannot adequately account for those in which the two families live far apart and have apparently had no previous contact with each other.[1] The informants' testimony may, however, be unreliable in other respects. They may show a type of paramnesia. The families concerned in these cases have nearly always met before the cases are investigated, and it is possible that, after they meet, they mingle their memories and credit the child subject of a case with having had more accurate knowledge about the deceased person whose life he claimed to remember than he really did have before the families met. This would not need to occur fraudulently; the informants might show varying degrees of forgetfulness and embellishment of which they remain innocently unaware. In a small number of cases - sixteen in all - a written record of what the child said about the previous life was made before anyone verified his statements. Distortions of memory could not account for these cases, although their occurrence does not necessarily justify confidence in all the others. Yet paramnesia on the part of informants does not seem a plausible explanation for the observation that similar features - such as a high incidence of violent death in the persons whose lives are remembered - frequently occur in widely separated cases. How can so many diverse informants, who are ignorant of each other, all have happened to say pretty much the same thing, unless they were reporting some widespread genuine phenomena? An obvious answer to this question is that (up to 1974) they were all - or nearly all - speaking to the same investigator (I.S.), who may have unconsciously guided them to make similar statements and to give similar answers.

The factors just mentioned generated a wish for a replication of I.S.'s investigations, and this led to the research on which the present report is based. S.P. investigated a series of cases independently of I.S. This report compares the results she obtained with data from another series investigated earlier by I.S. We call S.P.'s investigations a "partly independent" replication, because I.S. trained S.P. in his methods of investigation. Moreover, funds he controlled defrayed most of her expenses during her research. We shall consider later the extent to

which these factors may have biased S.P.'s investigations toward the results she obtained.

We wish to emphasize that the present paper is concerned primarily with the authenticity of the cases and only indirectly with the question of the evidence of paranormal processes in them.

METHODS

Subjects

The principal subjects were persons of northern India (usually children under the age of eleven) who claimed to remember details of a previous life. Newspaper reports furnished first information about some cases; private informants told us about most of them.

Among a larger number of cases than are here reported, some were rejected on the grounds of insufficiency of data for analysis. We set aside a few others because of grave doubts about the authenticity of the information given us. We are not aware of any other reasons for rejecting a case.

Ninety-five cases were selected for analysis. There were fifty-seven male and thirty-eight female subjects. Many of the subjects were still children under the age of eleven when their cases were investigated, but some were older and a few were adults. The median age of all subjects was 10.7 years.

The ninety-five cases were divided into two groups identified as "I.S. cases" (N = 50) and "S.P. cases" (N = 45). An "I.S. case" was defined as one for which I.S. had (with other interpreters and assistants) obtained the principal data before S.P. joined him as an assistant and interpreter. An "S.P. case" was defined as one for which S.P. obtained the principal data independently of I.S., that is, when he was not with her.[2] S.P., who speaks Hindi and Punjabi, needed no interpreter, but she was often assisted (when she was not with I.S.) by other Indian psychologists who had worked with him.

Techniques of Data Collection

The interview was the principal method used for collecting data. The interviewers gave the informants only the minimal guidance needed to obtain adequately comprehensive knowledge about

the case. Questions were used to fill gaps in essential information. With rare exceptions, we accepted only firsthand information; and we ourselves verified independently with the family of the deceased person concerned the statements attributed to the subject by informants for his side of the case. Written documents, such as medical reports and birth and death records, were studies whenever they seemed pertinent and were available. One of us has given elsewhere detailed accounts of the methods used in the investigation of these cases (Stevenson, 1966/1974, 1975).

The questioning of the informants elicited data on eighty-six variables. Not all of the variables, however, were suitable for comparison between the two series of cases. For nine variables the data were insufficient or judged unreliable. For example, data concerning cases of the "sex change" type were insufficient and data concerning birth order were judged unreliable. Although birth order is, in principle, an interesting and potentially important variable, we found that we had not been questioning the informants (usually the mother, sometimes the father) about it in a sufficiently detailed manner. We later realized that some informants had calculated the birth order by including stillborn infants and children who had died prior to the interviews, whereas others had considered only their living children. (We have since adopted uniform rules for recording data on this variable.)

Of the remaining seventy-seven variables, sixteen concerned demographic features, fifty-seven concerned the main features of the cases, and four concerned aspects of the investigation of the cases. All these variables were originally considered separately, but later some of the data on the main features of the cases were combined for convenience. Combining several variables closely related in content, such as different types of unusual behavior, reduced the variables concerning the main features of the cases from fifty-seven to thirty-six. No data concerning the demographic features or the investigation of the cases were combined.

All seventy-seven variables, however, were first analyzed to compare the two series, and those with significant differences were retained separately and not combined in any way. No variable, therefore, was omitted from the comparison between the two series on the basis of known or expected differences.

After all eliminations and combinations, there remained fifty-six variables suitable for a meaningful comparison between the two series.

Data Analysis

Fifty-six features of the cases were compared in the two series. These were divided into a group of sixteen items of demographic characteristics; a group of thirty-six items of the principal features of the cases, such as the frequency with which particular details, e.g., the name of the presumed previous personality and the mode of his death, figured in the subjects' statements; and four items concerning the investigation of the cases in the two series.

Appropriate statistical analyses (chi-square test or median test) were then applied to the differences between the data of I.S. cases and S.P. cases. Forty-five variables were recorded as either "present" or "absent", and they were therefore analyzed by the chi-square test. (Yates's correction was applied in all cases having a frequency of five or less in any cell.) The remaining eleven variables were scalar with a skewed distribution; therefore the comparison between the two series was made with the median test.

RESULTS

There were statistically significant differences between the data of the two series of cases with regard to twelve of the fifty-six variables compared. Of these twelve items, three occurred among the sixteen items of demographic features; six occurred among the thirty-six items concerned with the principal features of the cases; and three concerned the investigations in the two series of cases.

Comparisons of the Two Groups with Regard to Demographic Features

The two groups showed no significant differences with regard to the sex of the subjects investigated, their social and economic statuses, or their families' membership in religious groups. There were also no significant differences for these features between the concerned previous personalities of the cases in the two groups.

The subjects in the I.S. group lived more often in towns and cities, whereas those of the S.P. group lived more often in villages and small towns ($p < 0.01$). The concerned previous personalities in the I.S. group also lived more often in towns and cities compared with the corresponding persons in the S.P. group ($p < 0.01$).

The two groups showed a significant difference ($p < 0.01$) with regard to the education of the subjects' fathers. Those of the I.S. group had, on the whole, a higher education than those of the S.P. group. (We obtained insufficient information for a comparison of the educational level of the subjects' mothers.)

Comparisons of the Two Groups with Regard to the Main Features of the Cases

Age and Circumstances of the Subjects' Speaking about the Previous Lives. The two groups showed no significant differences with regard to the ages when the subjects first spoke about the previous lives. The median age was thirty-six months both for I.S. and S.P. cases. There was also no significant difference between them with regard to ages when the subjects stopped speaking spontaneously about the previous lives they claimed to remember.[3] This median age was seventy-eight months for I.S. cases and sixty-nine and a half months for S.P. cases. Nor were there significant differences in the circumstances that appeared to stimulate the subject's memories.

Principal Contents of the Subjects' Claimed Memories and Related Behavior. The two groups showed no significant differences with regard to the frequency with which the subjects included in their memories the name of the concerned previous personality and his mode of death. Seventy-six per cent of the subjects in I.S. cases mentioned the name, as compared with eighty-eight per cent of the subjects of S.P. cases. Seventy per cent of the subjects of I.S. cases recalled the mode of death in the previous life, and sixty-seven per cent of the subjects of the S.P. cases did so.

Unusual behavior (such as phobias) was noted with equal frequency in the two groups, but it was reported as persisting longer in the subjects of the I.S. group than in the subjects of the S.P. group ($p < 0.05$).

Both the distance between the subject's birthplace and the concerned previous personality's main residence and the distance between the subject's birthplace and the related previous personality's place of death were greater in the I.S. cases than in the S.P. cases ($p < 0.05$ for both comparisons). The median distances for these variables were respectively 45 and 46 kilometers in I.S. cases but 8 and 10 kilometers in S.P. cases.

The subjects in the I.S. group tended to remember previous lives in an economic status higher than their own; subjects in

the S.P. series did this less frequently, and the difference between the two series was significant ($p < 0.05$).

Informants reported evidence of extrasensory perception (other than the claimed memories of previous lives) on the part of the subjects more often in the I.S. group than in the S.P. group ($p < 0.01$). Observers also reported an "adult attitude"[4] more often for the subjects of the I.S. cases than for those of the S.P. cases ($p < 0.05$).

Prominent Characteristics of the Concerned Previous Personalities.

A person corresponding to the subject's statements about the previous life was satisfactorily identified in eighty per cent of the I.S. cases and in eighty-four per cent of the S.P. cases; the difference is not significant. Among all concerned previous personalities, the incidence of violent death was equally high (48%) in both groups.

The two groups did not differ significantly with regard to the concerned previous personalities' ages at death. The median age at death of these persons was thirty-one and one-half years for I.S. cases and forty-seven years for S.P. cases. The median interval between the death of the concerned previous personality and the subject's birth also did not differ significantly between the two series. It was nineteen months for I.S. cases and fourteen and one-half months for S.P. cases.

Connections between the Two Families Concerned before the Development of the Case. The two groups showed no significant differences with regard to biological, marital, social, or commercial relationships between the families before the development of the case that concerned them. Each group included some cases in which the two families concerned had had considerable contact before the case developed. There were seven (15.1%) "same family" cases in the I.S. series and six (13.6%) in the S.P. series. In addition, there were three (6.4%) cases in the I.S. series and five (11.1%) in the S.P. series that were placed within the category of "considerable contact". By this phrase we mean that the two families immediately concerned, although not actually related, were well acquainted through friendship or otherwise. In other cases the two families had had some minor contact before the case developed. In a small number of cases, the possibility of some contact - for example, because both families used the same market - could not be excluded. The existence of such a possibility, however, does not mean that any contact actually occurred. In both groups there

remained an important residue of cases in which the two families had had no acknowledged contact prior to the development of the case. Among the forty I.S. cases with a satisfactorily identified previous personality, the informants for twenty-two (55%) of the cases denied any association whatever with the other family concerned in the case. Among the thirty-eight S.P. cases also having a satisfactorily identified previous personality, the informants for fifteen (39.5%) of the cases denied any prior acquaintance with the other family. The difference between the two series in this variable was not significant.

Comparison of the Investigations of the Two Series of Cases

I.S. or his representatives arrived to investigate his cases within a shorter time after the meeting between the two families concerned than S.P. did for her cases. The median interval for I.S. cases was eighteen months, whereas for S.P. cases it was eighty-three and a half months; the difference was significant ($p < 0.05$).

There was no significant difference in the ages of the subjects of the two groups at the time the cases were first investigated; this median age was 9.3 years for subjects of I.S. cases and 12.5 years for those of S.P. cases. The subjects of I.S. cases were, however, significantly older than the subjects of S.P. cases at the time of the latest interviews, usually for follow-up studies; the median age of subjects of the I.S. cases at this time was 15.5 years and the corresponding median age for subjects of the S.P. cases was 12.9 years ($p < 0.05$).

The cases of the I.S. group were more thoroughly investigated[5] than those of the S.P. group ($p < 0.05$).

DISCUSSION

Interpretations of the Differences between the Two Series of Cases

Although the majority of the comparisons between the two series yielded no significant differences, some (twelve out of fifty-six) did. In considering various reasons for these differences between the two groups of cases, we were aware that some of them might have resulted from differences in the samples. It is also possible that informants for the cases observed and reported some features differently. In addition, other differences might be attributable to inconsistencies

between the investigators in collecting data; this would indicate the need for further improvement and standardization in methods.

I.S. obtained first information about many of his cases from newspaper reports; S.P. obtained more of her first information from informants for other cases. Cases in towns and cities are more likely to come to the attention of journalists and be reported in newspapers than are those occurring in villages. The residents of towns and cities are usually better educated than villagers and therefore perhaps more sensitive to unusual behavior patterns such as adult attitude and indications of possible extrasensory perception on the part of their children. The differences between the two series in these two variables may also have derived from different techniques of interviewing on our part. I.S. might have given more attention to these matters than did S.P. and therefore elicited more information about them; alternatively, he may have unconsciously "forced" affirmative responses more often than S.P.

There are several possible reasons for the difference in the time required for the investigator to reach the cases. Cases reported in newspapers are usually more current (when published) than those reported by private informants. Circumstances of I.S.'s early work in northern India, where he had a team of associates and interpreters ready to assist him, allowed him or his associates to reach the scene of a case fairly soon after they were notified about it. In contrast, S.P., who was based in Bangalore (southern India), did not have the same easy access to cases in northern India. Also, S.P. learned about the location of many of her cases from I.S. (who had been informed about their existence by his associates or other correspondents in India), and the time required in passing on information about new cases to S.P. sometimes caused another delay.

We think that three of the differences in the two series derive from the different periods during which the cases developed and were investigated. These are: the thoroughness of the investigations, the age of the subjects at the latest interview, and the persistence of unusual behavior on the part of the subjects. I.S. has made field investigations in India since 1961; S.P. has investigated cases only since 1974. Because I.S. has often returned to old cases for checking further details and for follow-up interviews, his cases were, on the whole, more thoroughly investigated than those studied by S.P.

Similarly, because I.S. was able to follow his cases over a longer period of time than S.P. has been able to follow hers (up to now), he had more opportunity to observe the persistence or cessation of unusual behavior on the part of the subjects of his cases than did S.P.

The I.S. group contains a larger number of "long distance" cases than the S.P. group; that is, there was a tendency for the subjects of this group to be born at a greater distance from the residence and place of death of their related previous personalities. Since, however, we regard geographic distance as one of the objectively verifiable[6] variables, we believe that the difference between the series in this variable represents a real difference in the samples studied rather than a possible bias in the recording of data on the part of one or the other investigator.

We have no explanation for why the subjects of I.S.'s cases remembered previous lives in superior economic circumstances more often than the subjects of S.P. cases.

Interpretations of the Similarities between the Two Series of Cases

Since the similarities in the two groups of cases are much more numerous than the differences and since, in our opinion, they occur in the more important features compared, we must consider possible interpretations for the reporting by S.P. of a group of cases that resemble closely in their main features those investigated by I.S.

Was S.P. a mere trainee and assistant of I.S. who did not dare obtain results that differed from his? We do not believe she was. S.P. had over five years of research experience before she joined I.S. as an interpreter and assistant. We deny that she was a compliant instrument striving to produce data similar to those that I.S. had obtained. More importantly, many of the data were of a kind that we call "objectively verifiable" (see note 6). Such data include dates, names of places and persons, and other information about which different recorders would not disagree. It would be virtually impossible for unconscious biases to influence the recording of data of these kinds. And finally, even if S.P. had wished (unconsciously) to make her data conform to those of I.S., she could have obtained almost no guidance from his published case reports and other writings (Stevenson, 1975, Stevenson et al., 1974), which contained only brief allusions to a few recurrent features in large numbers of cases.

Thus S.P. could not have known consciously about all the features of a typical Indian case suggestive of reincarnation before the analyses of the present report were made.

We cannot, however, exclude the importance of some subtle influence on S.P. by I.S., especially in the collection of less objective data. During the period of S.P.'s training, which coincided with the period when she was collecting data for this project, they often discussed cases that they were investigating together. (These "joint" cases are not reported here.) At such times, I.S. may have communicated to S.P. more than he realized about his own emerging ideas concerning the characteristics found in a typical case of this type in India. Yet we think it unlikely that such communications of still undeveloped impressions about the cases could account for all the correspondences we found in the two series of cases summarized in this paper. Furthermore, I.S. had not carried out a systematic analysis of cases that he had investigated in India. For most of the data compared in the two series, therefore, the results of the present analysis provided the first comprehensive information either of us had about recurrent characteristics in a large series of Indian cases.

We have considered also the possibilitythat the data given by informants to both I.S. and S.P. all derived from some oral tradition about the features of a typical case of the reincarnation type in India. The informants for one case, however, rarely know any details about another. Also, although some informants had rudimentary ideas about some features of what we would call typical cases, such as the high incidence of untimely deaths, most of them had nothing like a clear concept of the features that have emerged from these investigations. Now for the first time, as reports of these cases are being published and becoming more widely known, some persons are learning about recurrent features in the cases. But little information about such recurrent features has reached informants in India.

Concluding Remarks

If we are correct in the foregoing arguments, then it is possible that the uniformities we found in the two groups of cases occurred, first, because the informants' reports were accurate - in general, if not in all details - and second, because they were describing events related to some genuine, that is, natural phenomenon.

We consider reincarnation only one of several possible

interpretations for the cases that we call "cases of the reincarnation type" or "cases suggestive of reincarnation". We are not here directly concerned with the various paranormal interpretations of the cases, for which we refer readers to our earlier reviews of these (Pasricha and Stevenson, 1977, Stevenson, 1966/1974, 1977b). Such interpretations only become relevant when we have satisfied ourselves that the reports of at least some such cases are authentic. We believe the data of this paper supplement other evidence (published in case reports) indicating that the reports of some cases are authentic. Because of the limitations we have acknowledged in the comparison reported here, we regard its results as only a first, but nevertheless important, step toward the replication of the investigations of these cases that I.S. has earlier reported.

ABSTRACT

The authors compared results obtained in two series of "cases suggestive of reincarnation" in India that they investigated independently. The phrases "cases suggestive of reincarnation" and "cases of the reincarnation type" refer to persons (usually children) who claim to have lived before and to remember details of the claimed previous life. There were fifty cases in the first series and forty-five cases in the second, later investigated series. Fifty-six items suitable for analysis were divided into three groups. The first group contained sixteen items of demographic features of the persons concerned in the cases; the second contained thirty-six items of the main features of the cases; and four variables concerned different aspects of the investigation of the cases.

For forty-four variables, no significant differences between the data of the two groups occurred. Significant differences between the series occurred for three items of the demographic variables, six items of the main features of the cases, and three items concerning the investigations of the cases of the two series.

The authors believe that the second investigation constitutes a replication, although not an entirely independent one, of the earlier investigation. The similarities in the two groups of cases do not seem to derive from conformity on the part of the second investigator or on the part of informants for the two series. The authors think the resemblances provide additional evidence for the authenticity of the cases that were included in

these two series.

NOTES

1 This statement refers primarily to cases in India and Sri Lanka. In the cases of some other countries, such as Burma and Turkey, the two families concerned have frequently known each other before the development of the case, although often only slightly. In some cases (even in India) both persons concerned - the subject and the person whose life he remembers - belong to the same family.

2 In three of the S.P. cases, I.S. participated in the data collection to a small extent; but we do not think his activity in these cases affected the essential independence of S.P.'s investigation of them.

3 Careful questioning of the subjects and parents concerned in these cases often shows that the subjects stop speaking spontaneously about the previous lives before they have actually forgotten all the details apparently remembered. For practical purposes, the age of a child's no longer speaking spontaneously about the previous life he claimed to remember may be taken as the age of fading of his memories; but in individual cases wide departures from this general relationship may occur. One of us (I.S.) has discussed elsewhere (Stevenson, 1966/1974) the difficulties in assessing reports of actual fading of the memories or claims of preserving them.

4 The term "adult attitude" is used to designate a group of behaviors in which the child assumes attitudes (and often responsibilities) markedly beyond those of other children of his age. For example, a younger daughter in a family who recalls the previous life of a woman who was still raising young children when she died may assume mothering roles with other children, including her own older siblings. She may care for them in various unexpected ways and give them monitions in the manner of a mother looking after her children.

5 We evaluated "thoroughness of investigation" on a scale that took account of the number of informants interviewed for a case and the number of interviews we had with them. It also included an assessment (necessarily somewhat subjective) of the amount of information obtained for the case in relation to the total information theoretically obtainable for it.

6 By "objectively verifiable", we do not mean accurate. The question of accuracy, especially in dates, is a separate one; and we are far from satisfied with the accuracy of all the dates we recorded. By "objectively verifiable", we mean that different recorders of the same information could be expected to note down, for example, the same date, unless one or the other of them was deliberately falsifying records, which is not here in question. Other data, such as judgments about the social and economic status of the informants, are not objectively verifiable. (At least they are not so unless special instruments are used for recording data from which a grade for social or economic status is then derived; S.P. did, on most occasions, use such an instrument, but I.S. did not.) Two investigators who observe the same house with its contents may record quite divergent judgments about the owner's social and economic status. Even statements by parents about a child's age of first speaking coherently, or first speaking about a previous life, may vary between parents. And each parent may give different statements about these ages on different occasions of interviewing. We would not consider the recording of this type of information objectively verifiable.

ACKNOWLEDGEMENTS

This article is based on a paper presented at the 22nd annual convention of the Parapsychological Association, held at St. Mary's College, Moraga, California, August 15-18, 1979.

Much of the data of this paper derives from research work for a thesis submitted by one of us (S.P.) in fulfillment of the requirement for a doctoral dissertation at the National Institute of Mental Health and Neuro-Sciences, Bangalore University, Bangalore, India.

Thanks are due to the James S. McDonnell Foundation, the Bernstein Brothers Foundation, and the John E. Fetzer Foundation for the support of this investigation. Emily F. Williams contributed helpful suggestions for the improvement of the paper.

REFERENCES

Pasricha, S., and Stevenson, I. Three cases of the reincarnation type in India. Indian Journal of Psychiatry, 1977, 19, 36-42.

Stevenson, I. The evidence for survival from claimed memories of former incarnations. Journal of the American Society for Psychical Research, 1960, 54, 51-71, 95-117.

Stevenson, I. Twenty Cases Suggestive of Reincarnation (2nd Ed.). Charlottesville: University Press of Virginia, 1974. (First published in 1966.)

Stevenson, I. Cases of the Reincarnation Type. Vol. 1. Ten Cases in India. Charlottesville: University Press of Virginia, 1975.

Stevenson, I. Cases of the Reincarnation Type. Vol. 2. Ten Cases in Sri Lanka. Charlottesville: University Press of Virginia, 1977 (a).

Stevenson, I. Reincarnation: field studies and theoretical issues. In B. Wolman (Ed.). Handbook of Parapsychology. New York: Van Nostrand Reinhold, 1977 (b), pp. 631-663.

Stevenson, I., Prasad, J., Mehrotra, L.P., Rawat, K.S., The investigation of cases of the reincarnation type in India. Contributions to Asian Studies, 1974, 5, 36-49.

Satwant Pasricha
Ian Stevenson
Division of Parapsychology, Box 152
Department of Behavioral Medicine and Psychiatry
University of Virginia School of Medicine
Charlottesville, Virginia 22908
U.S.A.

O B I T U A R Y

J. Gaither Pratt 1910 - 1979

The moment this issue of the European Journal of Parapsychology was to leave my Editor's desk, the sad and shocking news of the death of Dr. J. Gaither Pratt arrived. Being one of the most distinguished parapsychologists and perhaps the most experienced of all experimental parapsychologists, Dr. Pratt was also a man of great integrity and a personal friend of long standing.

At the age of 69, J. Gaither Pratt passed away on November 3, 1979 at his home, near Charlottesville, Virginia. Surviving are his wife, Ruth Pratt, three sons, John, Vernon and Joseph, a daughter, Ellen, a brother, five sisters, and six grandchildren.

Born in 1910 as the fourth child of a farming couple in the Piedmont section of North Carolina, near Winston-Salem, Gaither Pratt was raised in a bustling family of twelve. At a very early age he indicated his wish to become a Methodist minister. After College he entered the Duke University School of Religion. His scientific attitude, however, caused him to change his career. In his book "Parapsychology: An Insider's View of ESP", Pratt states: "Not until the first semester of graduate study in the Duke University School of Religion did I finally pause long enough in the pursuit of this goal to realize that my mind was not suited to a profession in which the answers to the great questions regarding man and his relation to the universe are largely taken on faith ... In February of 1932, I switched courses and entered the Duke Department of Psychology" (p. 8). That same year he became Dr. J.B. Rhine's research assistent in experimental work in the area of ESP. Dr. Pratt took his M.A. in 1933 and obtained his Ph.D. in 1936, though surprisingly not on a thesis on ESP but on an investigation of the nature of learning in white rats. After his thesis work, he spent two years at Colombia University, working with Dr. Gardner Murphy in ESP research. How he finally decides to make parapsychology his profession, is related in his above quoted book: "The meeting at which I decided upon parapsychology as a profession took place during the second year in New York. Dr. Rhine, Dr. Murphy and I were in Dr. Rhine's hotel room ... We were discussing plans for the establishment of what we hoped

would become a permanent parapsychological laboratory at Duke. Would I be interested in coming back to Duke next fall as a full-time member of the research staff? ... There is no stop-watch record of the time I took to decide, but I do not remember having any difficulties. I cast my life for parapsychology" (p. 10).

As all readers of parapsychological journals certainly know, Dr. Pratt stayed on in Durham for more than a quarter of a centuary (1937 - 1963). At the Parapsychology Laboratory he made a great number of contributions, not least in the field of assessing paranormal research data statistically. Several of his experimental series are considered "classics" in parapsychological literature. His extensive studies of Pavel Stepanek belong to the important events in parapsychology during the sixties. Quite properly he received the 1970 McDougall Award (together with Dr. Jürgen Keil) for Distinguished Work in Parapsychology.

Dr. Pratt was probably the first to visualize the great potentials of using computers in parapsychological research in order to analyze data and monitor experiments. During his early years at the Division of Parapsychology at the University of Virginia, Charlottesville, he labored on the development of suitable soft-ware for computer-analysis of Psi-data. At the research center in Charlottesville, where he arrived in 1964 and where he was appointed Professor of Psychiatry, he became more and more interested in field-studies of poltergeist phenomena. He produced more than a hundred professional papers and articles, in addition to several books of which "Parapsychology: An Insider's View of ESP", ESP Research Today: A Study of Developments in Parapsychology since 1960" and "The Psychic Realm: What Can you Believe?" probably are the best known.

Dr. Pratt took part in the creation of the Parapsychological Association, served as its President and was elected council member several times. He also played an important role on the council of the American Society for Psychical Research, and, too, was President of the Board of the Psychical Research Foundation, Inc.

One should remember Dr. Pratt as an important ambassador of parapsychology. He made a great number of trips for research, visiting western European countries, the Soviet Union, Japan, India and Africa.

.--.

I shall always treasure the warm personal friendship that developed

between Dr. Pratt and myself. He was one of the persons instrumental in the channeling of my career in parapsychology.

During the sixties he paid many visits to Lund, often on his way to or from Prague. My wife and I always looked forward to having Gaither Pratt as our guest. We also were fortunate to have several opportunities to visit the Pratts at their scenically located farm, "High Fields", where they raised Angus cattle and grew their own vegetables. Dr. Pratt, a warm lover of nature and animals, was deeply concerned about ecological problems and man's abuse of raw materials. He was a very modest man who maintained through out his entire life a strong fascination for the miracles of mother nature. He was a kind and lively person. In a social setting he enjoyed punning games and was just as much fascinated by his own thoughts and experiences as those of others.

In forthcoming parapsychological meetings there will be a great vacuum after Dr. Pratt.

We all have good reasons to feel gratitude for his contributions to our field. Those of us who had the privilege to be his personal friends, will of course mourn him especially. We will also bear our deepest feelings of compassion with his courageous, very talented and profiled wife, Mrs. Ruth Pratt.

Martin Johnson

THE CATEGORIES OF PSI: THE CASE FOR RETENTION

John Beloff

University of Edinburgh

There are two creditable reasons that I can think of for applying Occam's razor so as to reduce the number of basic concepts in a given science. The first of these is that, other things being equal, the fewer the assumptions we require the more powerful and more elegant our theory. A classic instance of this sort was the elimination of the concept of an ether, which had played such a prominent part in the theorizing of physicists during the 19th century, with the advent of relativity theory. Of course, parsimony is not always successful. Einstein's lifelong attempt to go further and produce a unified field theory which would embrace both gravitation and electromagnetism was fruitless and today, despite all the talk about 'quarks', modern physics has still been unable to reduce either its plethora of so-called fundamental particles or its four distinct forces. Nevertheless, no one disputes that parsimony is a legitimate aim in physics or that the fewer the basic concepts and entities a scientist needs to posit the more advanced the state of his science.

The second reason for wishing to get rid of redundant categories stems not so much from a love of parsimony than from a dissatisfaction with some particular category and from a feeling that we would be better off without it. One could cite in this connection the concept of 'phlogiston' in 18th century chemistry about which controversy was especially acute when it became evident that one would have to assign it negative weight. With the discovery of oxygen and of the laws of combustion it was duly consigned to the scrap-heap and expunged from the vocabulary of science. Or, to take an example from the history of medicine, the concept of animal magnetism, the fluid of the mesmerists, eventually dropped out of circulation when the psychology of hypnosis and suggestibility became more clearly understood.

Note: Based on a paper delivered to the Parascience Conference, London, September 1979.

Now, the first of my two reasons would seem to have little relevance outside the domain of the advanced theoretical sciences. And, since I do not think that any of us are brazen enough to claim that parapsychology has reached this advanced stage, despite the worthy efforts of the so-called paraphysicists, the only reason we can give for wanting to get rid of this or that parapsychological concept is that we dislike it, or disapprove of it, whether on methodological, metaphysical or other grounds. But these grounds will have to be very cogent indeed if we are to escape the charge of acting capriciously. For the strength of parapsychology, so, at least, it has always seemed to me, lies precisely in its resolute empiricism, in its insistence - almost unique in science - on putting facts and observations above all else no matter how anamolous or how unexpected they may be in terms of our preconceptions or intuitions. And, once we sacrifice this virtue for the sake of some pet theory which we desire to promote, we run the risk of arbitrarily distorting the evidence, the first step very often in the creation of some new occultism.

Having now explained why I am suspicious of any move designed to eliminate one or more of the traditional categories of psi phenomena, I wish to say something about what these categories are and which, if any, are ripe for elimination. At one extreme there is a case for recognizing only one basic psi ability or psi process and treating all the familiar manifestations as just derivatives or variants of this one fundamental source. Such a move would derive some plausibility from the fact that, in almost every case, someone who has shown outstanding psi ability of one sort will be found to exhibit psi ability of other kinds as well. At the opposite extreme one could argue in favour of accepting the classification we have inherited even though it grew up partly as a result of historical accident and owes more than a little to spiritualist beliefs and practices. Consider, for example, the so-called 'out-of-body' experience, once better known as 'astral projection'. Is this a category in its own right or is it just a mode of ESP with occasional traces of PK? Or, consider the concept of the poltergeist, a time-honoured concept if ever there was one. Can this be assimilated to the concept of PK as this is known from laboratory experimentation? This, at any rate was the idea behind the expression 'RSPK' (Recurrent Spontaneous PK) which the American School of poltergeist

study, as represented by W.G. Roll, now regularly use. Yet it has been contested in this country by Alan Gauld who, in his new book on poltergeists (Gauld & Cornell 1979) repudiates the term on the grounds that we cannot yet be sure we are dealing with the same phenomenon as that which we study in the laboratory, that, indeed, we cannot even be sure that poltergeist phenomena do emanate from the so-called poltergeist agent and not, on occasion, from some discarnate agency. Of all the traditional categories PK is the most troublesome from this point of view. For the variety of known or claimed paraphysical phenomena, from psychic photography to teleportation, materialization and psychic healing, is so vast that we are bound to hesitate before lumping them all together under the same rubric.

At the same time, one does not want to retain too many trivial distinctions that are clearly no more than terminological conveniences. We would not want to call 'clairaudience' a different phenomenon from clairvoyance, we would say rather that it was a form of clairvoyance which used the auditory rather than the visual modality. Similarly, psychometry is best regarded, not as a phenomenon in its own right, but as a particular traditional method of inducing clairvoyance. I would say the same about dowsing despite its distinctive mystique, its specialized aims and its specialist practitioners. What, then, are the minimum set of concepts which are truly indispensable for any parapsychological science?

Obviously, one needs the basic distinction between the mental and the physical phenomena, between ESP, conceived of in its most general sense as the paranormal acquisition of information or knowledge, and PK, conceived of in its most general sense as the paranormal production of certain physical effects. For, even if we were to adopt an 'observational theory' of psi, according to which psi is always in the last resort a question of PK (Millar 1978), even then we still require the logical distinction between the informational and the physical aspects of the phenomenon. On the other hand, time and space ought not, I suggest, to enter parapsychological theorizing at the primary level since it looks more and more as if psi phenomena are, in the strict sence, space-time independent. Hence, whether the target in an ESP test pertains to the past, the present or the future may be regareded as a secondary consideration. It is noteworthy that retroactive PK, which has only recently come on the scene with the use of prerecorded digits, is not regarded as different in kind from ordinary contemporaneous PK.

From a pragmatic or heuristic standpoint, on the other hand, one may well be justified in retaining more than the minimum number of categories required conceptually. Thus, the implications, both theoretical and practical, of precognition are so immense that we need a special word to describe the extrasensory awareness of future events. For the same reasons we need to distinguish between telepathy, which presupposes a target-person or 'agent', and clairvoyance where the target is some external object or event. Equally, if nowadays less conspicuously, we need to distinguish between the case where the active component of the psi transaction is a living person as opposed to one where it is a deceased or disembodied entity. The term 'theta-psi' has recently been introduced by W.G. Roll to stand for the latter possibility. However, all these nice distinctions which are part of our parapsychological legacy have, over the years, been challenged, for one reason or another, by reformers who have urged us to reconstruct our conceptual framework in the interests of parsimony. In this paper I shall consider four particular categories which have been attacked as flawed concepts which deserve to be jettisoned. They are, in the order in which I shall discuss them, (a) precognition (b) telepathy (c) clairvoyance and (d) theta-psi.

Precognition, though it has seldom bothered the practising parapsychologist, has proved a serious stumbling-block for theorists including not a few eminent philosophers, notably C.D. Broad and Antony Flew. Thus it has been said that, if taken literally, precognition is a logical absurdity inasmuch as it implies that an effect can precede its cause or, worse still, that an event which has not yet occurred can exert an influence on the present. However, the argument that an effect cannot precede its cause has since been shown to rest on an erroneous view of the logic of causation (Chisholm & Taylor 1960, Brier 1974). Backward causation may indeed be paranormal but it does not violate any known logical principles. As for the idea which worried Broad, that since future events do not exist they cannot be the causes of anything, this simply begs the whole question about the nature of time and of reality and to use it as an argument against the possibility of precognition is surely a case of putting the cart before the horse. This is not to deny that there are some quite severe paradoxes associated with the notion of knowing a future event. There is, to start with, the notorious 'intervention paradox' - how can a precognition serve at once as a warning and as a prediction, as seems to be demanded? Then there is the fatalist paradox - if we could know ahead of time what we were going to do in a given situation how could we then <u>decide</u> what to do, or would we then be fatalistically

constrained to act as foreseen? Nevertheless, such paradoxes can to some extent be met either by assuming certain boundary conditions within which the phenomenon must take place or by introducing a different model of the universe, for example one which would allow for potential as well as actual futures, but the point is that all paranormal phenomena are inevitably paradoxical to a greater or lesser degree and we must either try to live with them or else reject all parapsychological phenomena for this reason.

Since there are, so far as I can tell, no insuperable a priori objections to the concept of precognition we do not need to dwell here on the various proposals that have been made to circumvent it. We may recall, perhaps, Adrian Dobbs' ingenious idea that events might carry with them, so to speak, certain 'objective probabilities' of which we could become aware by extrasensory means and which we would then project onto the future(Dobbs 1965). It was even seriously suggested at one time that what we have been calling precognition might turn out to be our capacity to use PK in order to bring about (unwittingly) the fulfilment of our predictions (Roll 1961)! We must sincerely hope that this last suggestion, at any rate, is unwarranted if only because so many predictions concern disasters! I mention these far-fetched speculations because there is, it seems to me, a salutary lesson to be learnt from them namely that the attempt to obviate some traditional category of psi may involve us in a vast amount of wasted ingenuity.

J.B. Rhine has again and again in recent years (Rhine 1974) castigated telepathy as an 'untestable hypothesis'. He points out that, if we define telepathy as 'mind-to-mind communication', we are never in a position to know whether such communication has taken place or not. For, clearly, the only minds we can experiment with are those associated with a living brain and, once we bring the brain into the picture, we have a potential target for clairvoyant ascertainment even in the absence of any target object or symbol as would be the case in a so-called pure telepathy test. Hence, whatever information the subject gains in a test of telepathy could just as well be credited to his clairvoyance making the former redundant. Now there are two things that are wrong with this argument. First, even if we assume that a given thought is somehow encoded in a person's brain - an assumption which would be vigorously contested by some contemporary philosophers (Braude 1978) but which I shall assume here for the sake of argument - there is no reason to think that the way in which it is encoded in the brain of the agent would correspond to

the way in which it would be encoded in the brain of the percipient, if only because each person's brain is the product of a unique series of life experiences. Hence, even if the percipient enjoyed perfect clairvoyant powers and could scan the agent's brain in as much detail and with as much accuracy as he pleased, he would still be no nearer to interpreting the relevant target message. The naive view of telepathy as a direct transfer of a thought from one brain to another is thus a patent fallacy. But, even if we overlook these philosophical objections, a much more straightforward objection to Rhine's proposal is the one pointed out by R.H. Thouless (Thouless 1974), namely that to define telepathy as mind-to-mind communication is not only unnecessary but it gratuitously ensures that it will be operationally untestable. But all that is required to make telepathy meaningful is the assumption that the presence of an agent in some way affects the performance of the percipient and this proposition is eminently testable even if it should eventually turn out to be untrue. What makes Rhine's attempt to eliminate telepathy particularly unfortunate is that it happens to be the most widely accepted and intuitively plausible of all parapsychological concepts.

As compared with telepathy, the concept of clairvoyance is, admittedly, opaque (cf. Mundle 1967). Many who can conceive of there being some kind of a resonance between two sentient beings baulk at the idea that we can have direct access to the external world without the mediation of our senses. There just seems to be no ready analogy that we can draw upon to make such an idea intelligible. For example, clairvoyance implies that we could make colour discriminations even when the coloured objects are concealed in a light-proof container or that we could pick out the relevant target information even when, as in a 'down-through' test, this is concealed among a host of irrelevant information. Of course, these misgivings would have no force if we could be sure that we have empirical evidence for clairvoyance but, in fact, the number of experiments in the literature specifically designed to demonstrate pure clairvoyance (i.e. to exclude alternative interpretations in terms of telepathy, precognition or some combination of the two) is surprisingly small and the results are not all that impressive. For these reasons we should consider sympathetically the recent efforts of Peter Maddock to dispose of clairvoyance by invoking the possibility of precognitive telepathy as a counter-explanation (Maddock 1973). Briefly what Maddock is suggesting is that either the percipient is precognising the feedback which he will subsequently

receive or, where no such feedback is forthcoming, the experimenter serves unwittingly as the agent and the percipient precognizes telepathically what the experimenter sees when he looks at the results.

Is clairvoyance, then, the phlogiston of parapsychological theory? Unlike Rhine's objection to telepathy, which is a metaphysical muddle of his own making, Maddock's objection to clairvoyance stems from a genuine conceptual difficulty and his theory has the important merit of being falsifiable. For he agrees that, if we could demonstrate pure clairvoyance to our general satisfaction his theory would collapse. In the meanwhile, however, it is not good enough for Maddock or anyone else to defend this position by appealing merely to the weakness of the existing evidence for pure clairvoyance. For the fact is that there is a mass of evidence in the literature which, even if it does not exclude all alternative explanations, certainly suggests that clairvoyance is a fact, at least in the sense that the alternative explanations can be sustained only by the most circuitous reasoning. What is needed, therefore, if Maddock is to make his case is some positive evidence. Thus we would want to know whether a good ESP subject who is capable of scoring significantly when the conditions permit the use of some form of telepathy or precognition would invariably fail to score when conditions permitted only the use of clairvoyance. If I had access to a good ESP subject I think I would make such a project my top priority. Pending a resolution of this question I suggest that the rest of us would do well to suspend judgment.

We come, finally, to the last of our four disputed concepts, that which I am calling 'theta-psi'. Frederic Myers hoped that by establishing the reality of telepathy psychical researchers would render more plausible the idea of post-mortem survival inasmuch as telepathy between the living could then serve as a model for the communication between the living and the dead. J.B. Rhine, on the contrary, argued that, since parapsychology has shown that there are no known limits to the scope of ESP, ostensible communications between the living and the dead are more parsimoniously explained as due to the ESP abilities of the medium or sensitive through whom the messages are transmitted. Hence, according to Rhine, parapsychology has no further need of the PMS (post-mortem survival) hypothesis and it can be dismissed as just another untestable and time-wasting concept (Rhine 1975). This view is sometimes referred to as the 'Super-ESP' hypothesis (Gauld 1961).

Now I am sure we should all dearly love to know whether survival is a fact or a fiction, whether there are any discarnate entities in the universe and, if so, whether when we die we are destined to become such. For my part, though I find some of the survivalist evidence impressive as far as it goes, I do not think it obliges us to believe in the reality of theta-psi, there are other possibilities that must be discounted before we can accept such a momentous claim. However, it strikes me as nothing short of intellectual frivolity to pretend that there is simply no case to answer. Only the most naive view of science would sanction the idea that what cannot be proved must be rejected. And if we were to adopt such a view the logical conclusion would be that all psi phenomena must be rejected since none has been proved to the satisfaction of the scientific community. But the fact is that, outside of logic and mathematics, nothing can be proved in the strict sense, the most that science can demand is that a given conjecture be made plausible if it can be shown that it makes better sense of the evidence than any alternative conjectures. The correct question, therefore, concerning theta-psi is whether it makes better sense than any of the alternative explanations for a wide range of phenomena which include, hauntings, mediumistic communications etc. Our problem is to decide whether this is so or not, it is no part of our business to declare in advance that any interpretation is preferable to what is, after all, the simplest interpretation of the facts, namely survival. In conclusion, then, let us be very wary before we embark on any drastic tidying-up operation in parapsychology, we simply do not know enough about the phenomena we study to impose on them our personal preconceptions.Occam's razor must be wielded at all times with due care and delicacy if it is not to prove a two-edged weapon, for it can be used with equal facility by sceptics who seek to eliminate the whole of parapsychology.

REFERENCES

Braude, S. Telepathy. Nous 12, 1978, 267-301.

Brier, B. Precognition and the Philosophy of Science: An Essay in Backward Causation. New York: Humanities Press 1974.

Chisholm, R.M. & Taylor, R. Making Things to have Happened. Analysis 20, 1960, 73-78.

Dobbs, H.A.C. Time and ESP. Proc. S.P.R. 54, 1965, 249-361.

Gauld, A. The Super-ESP Hypothesis. Proc. S.P.R., 53, 1961, 226-246.

Gauld, A. & Cornell, A.D. Poltergeists. London: Routledge 1979.

Maddock, P. The 'Straight Bit' Systematization Scheme for the Mental Phenomena of Psi. Parascience Proceedings 1973-1977, Vol. 1, 23-37.

Millar, B. The Observational Theories: A Primer. European J. Parapsych. 2, 1978, 304-332.

Mundle, C.W.K. The Explanation of ESP. Internat. J. Parapsych., 7, Summer 1965 (reprinted in J.R. Smythies (Ed.) Science and ESP. London: Routledge 1967).

Rhine, J.B. Telepathy and Ohter Untestable Hypotheses. J. Parapsych. 38, 1974, 137-153.

Rhine, J.B. Comments: Psi Methods Reexamined, Sect II: Are we asking the right questions? J. Parapsych. 39, 1975, 267-301.

Roll, W.G. The Problem of Precognition. J.S.P.R., 41, 1961, 226-253.

Thouless, R.H. Letter. J. Parapsych. 38, 1974, 421-422.

John Beloff
Department of Psychology
University of Edinburgh
George Square, Edinburgh EH8 9JZ
Scotland

THE DISTRIBUTION OF PSI

Brian Millar

University of Utrecht

APOLOGIA

This paper critically discusses the common assumption that psi ability is widely distributed. The primary task here is the preliminary one of deciding what questions should be asked of nature and how she may be queried to best effect. There is no pretence of providing a conclusive solution to the points raised: on the contrary, only further work, both in detailed and extended analysis of the extant literature and in new experiments can be expected to provide final answers.

In one respect, however, this paper should perhaps be nailed to the door of the old Duke laboratory at dead of night. It is the author's view that the results of most positive experiments do not depend on the subjects at all, but on the experimenter. While strong evidence exists that experimenter-psi is sometimes involved, the task of this discussion is not that of reviewing the literature with respect to the strength or ubiquity of the effect, but instead it explores the consequences of such a view for the distribution of psi and thus for parapsychology in general.

INTRODUCTION

There are two opposed views on the distribution of psi. The first is that psi ability is rare but when it is present it is quite strong. This will be referred to as the psi-star hypothesis. Historically this idea has been expressed in the West in terms of Saints and witches. Furthermore much of the early work of psychical research was performed with specially gifted persons, generally mediums.

The second opinion is that most people have some psi which can be coaxed out under appropriate conditions, but which is usually quite weak. This will be called the democratic psi hypothesis. Some support for this comes from spontaneous cases. It became the prevailing view with the Rhine revolution of the 1930 s: after an initial period

in which the traditional policy of experimenting with high-scoring subjects was pursued, interest shifted rather abruptly to testing groups of unselected subjects (Pratt, 1977, p 74).

A belief in some generally distributed psi ability in addition to the strong powers of special individuals exists in many societies. Particularly in unspecialised societies in which each individual must be his own butcher or tailor he must also act as his own psychic when the need arises. It is perhaps of some sociological interest that the hypothesis of widely distributed psi found favour in the America of the '30 s with its aggressively democratic outlook.

In terms of experimental work the idea of widely-distributed psi is an optimistic one. On the psi-star hypothesis it is necessary first of all to find a psychic: on the democratic view of psi, however, any unselected group of people can be used to answer questions about paranormal effects.

One of the turning points in the history of parapsychology is the sheep/goat work of Gertrude Schmeidler. The major experiment reported together with McConnell (1958) used unselected classes of students. A total of 1157 subjects were tested with ESP cards giving 10,035 runs. The sheep (believers) scored above chance expectation while the goats scored below. The absolute size of the deviation was small (+ 0.10 and - 0.07 of a guess respectively per run of 25 trials), but because of the number of trials the difference between sheep and goats was unquestionably significant, with a probability of 3×10^{-5}. This seemed to set the seal upon the democratic psi hypothesis. Almost all parapsychological experiments since have tacitly assumed this model of psi distribution.

In addition to this practical work Schmeidler and McConnell devoted a chapter (How Common is ESP?) to discussion of the subject matter of this paper. At that time (1958) they stated that the evidence for which model was correct was inconclusive. Unhappily no substantial advance has been made on the issue since then. The empirical distribution presented by Schmeidler and McConnell is one of the few available in the literature. It seems almost inconceivable that such a vitally important practical question should be almost ignored by parapsychologists in favour of a theoretical predilection.

Schmeidler and McConnell's contribution, however, was by no means limited to the attempt to make an unequivocal deduction. On the contrary, on finding this attempt in vain they went the other mile and asked what the extant data suggested. Their conclusion was that "The evidence, from this research and from other sources, is .... consistent with the hypothesis that psi ability is widespread in the populations that have been tested.". They go on to comment on the

gifted subjects found by other experimenters "..... there is evidently a higher level of such ability possessed by only a few .... a level which is so much higher than the normal one that .... it is convenient to think of it as a different function.". In short they suggest that there are two parts to the distribution curve, a small scoring rate component due to a substantial portion of the population and a special subjects component, very much rarer, but scoring at extremely high rates. This compromise seems a very reasonable one on a descriptive level. On the conceptual plane, however, the ascription of the lower rate component to subject psi will be shown to be, at best, questionable.

THE DESCRIPTION OF PSI DISTRIBUTION

Frequently the view is expressed that psi is normally distributed. We may perhaps ignore the fact that there is probably no psychological variable known which is precisely normal (Dorfman, 1978, p 1180). In this context presumably what is meant is that the distribution curve rises smoothly to a peak and then drops away again approximately symmetrically. For this to be a testable statement we need to have some way of measuring psi. Should we perhaps use the psi-coefficient or Schmidt's theta, or would Walker's omega be best? (Timm, 1973; Schmidt, 1975; Walker, 1975). In practice we do not know which, if any, stays the same with changing p-values of the targets. It therefore seems best to avoid making an arbitrary decision and to work directly (at least initially) with the empirically observed distributions of scoring.

Consider recruiting several hundred experimenters who go out and test everybody they can with 9 runs of ESP cards under GESP conditions. Each subject tested then has a mean run score: the frequency distribution (or corresponding empirical probabilities) of these can be graphed. Figure 1 a, b and c show respectively the (psi-hitting) distributions that might be obtained if the population consisted solely of a homogeneous group of subjects with zero, a little and a lot of psi. While it is improbable that the real situation is composed of a mixture of a few homogeneous groups, as will be illustrated here, this clearly reveals the basic principles. It seems likely that any real population will generally contain a majority of subjects who show no psi under the conditions of the experiment (curve a). This may be mixed with b to produce figure 2 and with c to give figure 3, or indeed with both resulting in the distribution of figure 4. The dashed line indicates binomial (chance) expectation. In both figures 2 and 3 the curve is drawn as clearly bimodal: in practice however in figure 2, where the peaks of the curves are close together a bimodality will likely not be visible and the effect of psi-scoring will simply be to shift the curve somewhat to the right. The curve (fig. 2) represents the democratic psi distribution: many of the people tested show psi but at a rather

low level. Figure 3 represents the psi-star distribution: for clarity the height of the scoring peak is exaggerated. Here few people display psi but when they do it is at a rather high level. Figure 4 is the composite "Schmeidler-McConnell distribution". The empirical curves would specify the psi-scoring distribution under the particular conditions of the experiment and could be compared with theoretical ideas.

IS THERE A SINGLE PSI DISTRIBUTION?

If such empirical distributions were available, a large number of interesting modifications could be tried. For example, the experiment could be repeated with some allegedly psi-enhancing technique. Enhanced scoring means that either the non-chance peaks get bigger (more people score) and/or they shift to the right (the same number of people score but they do so at a higher rate). If the curve were of the Schmeidler-McConnell type (fig. 4) the possibility exists that the two scoring peaks might react in disparate ways. As another possible modification suppose we simply perform a different number of runs. A common simplifying assumption is that scoring rate is constant for a given subject. This is not generally true: declines are ubiquitous. Perhaps, for example, some subjects have only a certain amount of information to spend; thus they obtain the same significance level regardless of how many runs they do: the scoring rate will however decline with the number of runs (cf. Walker, 1976, p 49-50). Another simple comparison would be between ESP and PK.

To talk of such experiments is however, at the moment, somewhat utopian. Hardly any empirical data exist: some of these will be examined later. One point which it is important to make is that these curves may contain useful information about the psi process, information which is usually thrown away when attention is paid only to the mean and variance. It is only fair to add, however, that such work can normally be done only with a much larger data-base than is usually the case in parapsychological work: this has certainly contributed to the general experimental malaise.

It might be that the curves obtained under different conditions are quite chaotic and there is no relationship between them. Under these circumstances the distribution of scoring is restricted to a highly specific set of experimental conditions.

If, however, some degree of cross-condition predictability exists it might arise in two different ways. Firstly, subjects who score high under one set of conditions might also tend to score high under other conditions etc. This common factor could be regarded as a measure of ability. Most parapsychologists hold such a view. It is to be expected, however, that a rank correlation between different

Figure 1. (Psi-hitting) Scoring distributions for (a) no psi, (b) a little, (c) a lot

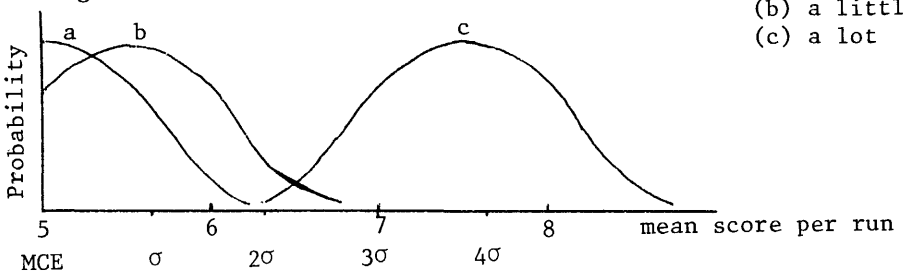


Figure 2. The democratic psi distribution

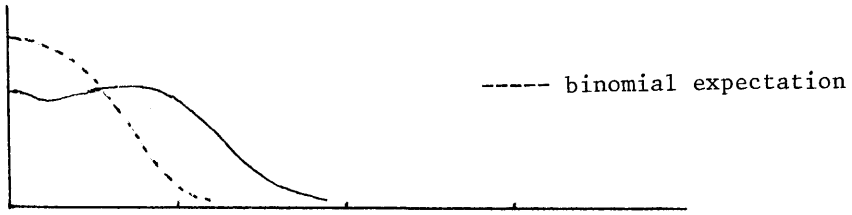


Figure 3. The psi-star distribution

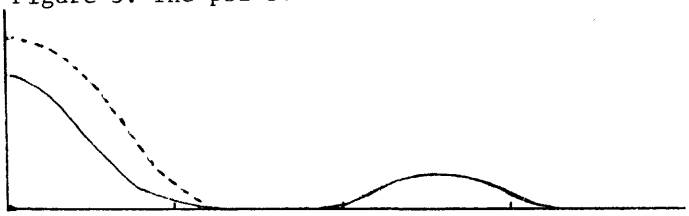
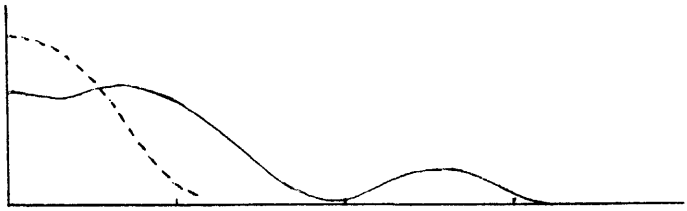


Figure 4. The Schmeidler-McConnell distribution



conditions would not be perfect but would be attenuated by individual differences. Secondly, it is possible that such intrasubject differences are predominant and although the features of one distribution might be predicted from another, the scoring would be due to different subjects in the two cases. In such circumstances no simple measure of psi ability can be defined.

POTENTIAL AND ATTAINMENT

It is general in psychology to draw a distinction between potential and attainment. Roughly, potential is the maximum possible attainment for a given person. There may be certain conditions in which a subject can score at a much higher rate than under other circumstances. How important this is in practice must be determined by experiment. In principle one could test a given population in some standard way and then in such a way that each subject works under his optimum conditions. This operational definition of potential, however, is a rather fuzzy one because we currently have little knowledge of how to set up the required optimum conditions. In practice the experiment would reduce to comparing scoring distributions under a number of different conditions. The question of interest is how and to what extent such manipulations would affect the distribution.

The idea of potential, as used here, requires some stability in scoring under the specified near-optimum conditions. A one-off situation can allow one at best to say "His potential was ....". Similarly, if for instance it were to transpire that all subjects very infrequently give exceptionally high scores, but at random with respect to any specifiable conditions then such a measurement of potential would have little, if any, predictive value.

WHAT POPULATION?

Sampling theory specifies that each subject unit used in an experiment should be sampled at random from the population of interest so that inferences may be made about that population. In practice, however, this is usually impracticable. The exigencies of life dictate the selection of the individuals tested and the population to which the results can be generalised is therefore to some degree fuzzy. Major characteristics are almost always clear, for example one might test schoolchildren, gerbils or Australian Aborigines. The most commonly used subjects in parapsychology, however, are university undergraduates. It has been said that experimental psychology is largely the study of psychology students.

Perhaps the closest approach to the requirements of sampling theory in common use is the employment of classes of students. In much work,

however, subjects are semi-selected by, for example, replying to newspaper advertisements. Even greater selection is made in some cases where successful subjects are re-cycled in experiment after experiment. Some of the altered states work, for example, uses this kind of subject selection. In such cases the attempt to generalise to a population is rather dubious.

For the purpose of constructing population distributions greater attention should be paid to subject selection than is usually necessary in parapsychological work.

WALKER'S WORK ON ESP DISTRIBUTION

Walker (1975, p 16-20) has tried to calculate from his theory the distribution of ESP scoring to be expected in the population. His theory is one of the observational type (Millar, 1978). In these theories all psi is basically time displacement PK. The psi source is triggered by the feedback of results at a later time and this triggering affects the random generator at the earlier time the target was generated. In the case of ESP the RNG is something in the subject's head at the time he made his guesses. The idea underlying Walker's treatment is that the brain is not a terribly good random generator. The observational theories in general assert that only pure chance events, ultimately derived from Heisenberg uncertainty on a quantum level, can be affected by PK (Millar, 1978). In most real systems, however, the bulk of the output is deterministic, while only a little is pure chance. The presence of the deterministic component, which cannot be affected by PK, would be expected to result in a lower PK-scoring rate than would a pure-chance RNG. Walker argues that the brain similarly contains a small proportion of noise upon which PK can operate and considers that this dilution is the major limiting factor in ESP ability. Further, he considers that chance variations between different people, in the way the counter-dilution strategy is applied, are the determiners of the ESP distribution.

In Walker's neurophysiological theory both the part that can be affected by PK (W) and the part that can not (C) are dependent on postulated quantum mechanical effects in the brain. Walker has calculated these data rates and it turns out that something like one bit in 10,000 (W/C) can be determined by PK influence.

Walker has to assume some model by which the brain constructs its guess from the two different kinds of noise. Unfortunately he does not spell this out in detail, but it appears to be something like the following. Before he starts to guess, a subject selects some method of classifying a subset of the bits which pass through his consciousness in the inter-trial interval. The selection of a classification procedure is the step which results in individual

differences, as will be seen later. The bits are used to add counts to a number of "bins", each of which represents one of the equiprobable targets.

For simplicity the principle will be illustrated with only two target alternatives: in this case each bit adds a count to one or other of the two bins. One might, for example, increment the bins in terms of the states of some progression of brain areas. It is assumed that the classification procedure, once decided, is adhered to for all trials. The origin of each such count would not be known to the subject; however it might for example turn out that bin one contains on each trial one count from data representing the itchiness of his big toe, a second count relevant to his breakfast that morning etc. Similarly a count might be derived from the W-channel and this correctly represents the target: such a count would be classified into the correct bin. Since, in the absence of learning, the subject has no knowledge of which counts are due to PK and which to non-psychic factors, the best he can do is to perform a majority vote. (It is tacitly assumed that MV here follows the normal statistics of signal enhancement.) On average, in the example, the correct bin (one) has one more count than the other (all the contents of which are randomly related to the target) and one will therefore become the slightly more probable guess alternative.

The assumption of consistency means that if a subject has happened to select a classification scheme in which, say three W-derived counts are classified into bin one (correctly) on the first trial, then he will continue doing this on subsequent trials. Consequently, the whole population will consist of people who regularly obtain 0, 1, 2, 3 .... correct counts per trial, in addition to irrelevant counts.

The relative frequencies of successive levels of ability is determined simply by the binomial expansion which gives the probabilities of choosing by chance, in the initial selection of classification procedure, successively 0, 1, 2, 3, ... PK-determined counts. The total number of bits (n) from which these are derived is determined not from within the theory but from the empirical proportion of the population showing no ESP. For convenience of computation the number n is assumed constant for all subjects. The scoring rates corresponding to the successive levels of ability are determined by the excess of 0, 1, 2, 3 .... counts in the correct bin, in the presence of all the other spurious counts distributed at random.

Using this treatment it is found that using ESP cards 9.5% of the population has one W-count per trial and scores around 5.6 hits per run; those with 2 counts (0.5%) score at 6.1; 3 counts (0.01%) at

6.6 etc. There are two interesting points to note here (1) according to Walker's empirical data 90% of the population does not score at all (though see later discussion) (2) the scorers come in a discrete series, rather than a continuum. Mathematically the discreteness arises from the numerical values of W/C and n. With other parameters the distribution may become (virtually) continuous.

One can criticise Walker's derivation on grounds of the arbitrariness of his model; however it is clearly necessary to assume something in order to obtain any result. More seriously, however, there is one certain and one probable error in the derivation.

Firstly, from the integral for genuine recognition probability (his equation 46) it seems that Walker assumes that one or another bin is taken as the guess if it exceeds MCE: in fact the bin taken as guess will be that which contains the most counts. The correct solution can be obtained from signal detection theory. In the multiple alternative forced choice guessing procedure of this theory, as in the present situation, one bin contains signal plus noise and the others only noise (McNicol, 1972, p 73). The distribution of the maximum noise is required: this is given by Tippett (1925). Hence the difference distribution between the signal bin and the maximum noise bin may be obtained and the genuine recognition probability can be calculated. Walker's result over-estimates this.

The second problem is with the extension of the model (illustrated above for binary choice) to the multiple choice situation. Walker uses every bit to increment one bin. This is perfectly admissible for the noise bits: it does not matter that each is constrained by the initially chosen classification scheme to increment only one of two rather than any one of the possible alternative bins. It does matter, however, for the signal bits. One possibility is that some W-bits will increment the wrong bins, since the right one is not one of the two allowed alternatives on that trial. The mathematics becomes complicated in this case. The other possibility is that a W-bit will always find some way to increment the correct bin; but in this case it would seem that a bit from the W-channel can be used to convey more than a single bit of information. This is the alternative implied by Walker's equations, though it may be meant simply as an approximation to avoid the mathematical problems of a more exact treatment.

PROPOSED MODIFICATION AND EXTENSION OF WALKER'S BASIC IDEA

The above technical criticisms are relatively unimportant in comparison with Walker's basic idea. Is ESP scoring primarily limited by the inability of the brain to function as a good random generator? In that case what about PK? If a pure chance RNG were used for a PK test with a successful ESP subject, then the PK scoring

rate on Walker's above model alone ought to be some orders of magnitude higher than for ESP, furthermore the rate should be the same for everybody. This is not true in practice: indeed, scoring is likely the same regardless of whether ESP or PK is involved, if the conditions are psychologically identical (e.g. Schmidt and Pantas, 1972). Walker has not explicitly extended his treatment to PK but his mention of ESP "decoding" and PK "encoding" (Walker, 1975, p 36) clearly suggests that he has in mind some similar treatment for PK.

If the scoring rates for ESP and PK under psychologically identical conditions are the same, however, this suggests that there are not two different processes involved but one, common to both kinds of psi. According to the observational theories, the step common to all kinds of psi is that of feedback. While it is, of course, possible to assert that psi sources are by their very nature of different strengths, this is not very illuminating. Another, and more interesting possibility, is that the psi source is only slightly affected by the feedback and to different extents for different subjects.

The important modification introduced here is consonant with the general picture of the observational theories: the things of importance happen at feedback and not, at the time of guessing for ESP, or target generation for PK.

Most work on ESP has centred upon the time the subject makes his guesses and has neglected the later feedback. There is empirical evidence that the state of the subject at the time of guessing has some influence on the result, however Stanford (1978, p 207) has suggested explicitly that the effect of many of the apparantly important psychological factors is simply to render the subject's brain at the time of guessing a better of worse random generator. He suggests, for example, that rational, sequential and contextual constraints make the brain a worse RNG and thus less responsive to PK influence, whereas in free-floating states like that of hypnagogic reverie it becomes a better random generator and hence can be more readily influenced by PK.

At first sight there seems to be a contradiction between the above emphasis on the pivotal role of the events at feedback and the empirically determined importance of the state of the subject at the time of guessing. However, these are not incompatible: it might be that the primary limitations on psi scoring are set at feedback, but the result of the psi source firing would be limited (perhaps in an all-or-nothing way) by the quality of the RNG on which it was working. In this case, the observed equality of scoring rates in psychologically equivalent ESP and PK tasks suggests that in the cases tested the

subjects's brains were acting as rather good random generators. This might, however, not be true in general: PK ability would therefore be expected to be more widely distributed than ESP. J.B. Rhine (1946, p 98) has indeed expressed just this view, judging by his own experience. However, there is a lack of experimental evidence bearing on the question.

In summary, the model proposed is that the major limitation on psi ability lies in how much effect the feedback information has in triggering the psi source. This may be described in information theory terms: the feedback signal is heavily perturbed inside the organism by extraneous noise on the channel leading to the psi source. The feedback has a different degree of effect on the psi source for different people. Thus, it is possible that the distribution of psi ability may ultimately be understood in terms of the normal neuronal circuitry between the sense receptor used by the feedback and the psi source. One complication arises when the RNG operated upon contains deterministic noise in addition to pure-chance noise. In ESP the brain of the subject may be such an imperfect RNG, the quality of which depends on his state at the time. Therefore ESP scoring may be less frequently observed (and/or less in magnitude) than PK scoring.

The main event in psi is characterized above as a sensory input signal causing firing in some normally unrelated brain area. The analogy which immediately springs to mind is that of synaesthesia in which the input of one sense is experienced in another as if via crossed lines; for example the sound of middle C might be accompanied by a green flash. This and the accompanying information theory concept suggest a number of novel possibilities.

Walker's suggestion that high-scoring subjects might come in a discrete series of scoring rates seems also to be worth at least some attention. If evidence for such an effect were to be found it could be of considerable theoretical value. A plot was made of the rates of the high-scoring subjects reported by Rhine (1964) but it revealed no evidence of scoring clusters. However, such a crude procedure is unlikely to reveal banding. It would be interesting to examine the run-score distributions within the data of high-scoring subjects. In particular, such a subject in decline might be expected to give a distribution consisting of the series of successively lower bands.

WHO IS THE SUBJECT?

In order to determine the scoring distribution of a population of subjects it is necessary to know who the subjects are. There is usually no problem in psychology: for example, in a sensory perception

experiment the subject is the person to whom the stimulus is presented and who makes the response. Making the response is rather less important if the stimulus is known to someone else, such as the experimenter. In such a case the possibility exists that the experimenter may subtly cue the subject via sensory means so that his responses may not be due to the sensory manipulation which the experiment is designed to explore. In parapsychology just who makes the responses is even less important, since these may be affected not merely by sensory means but also via psi. Consequently, the only way one could be sure in a psi experiment that the result was due to the subject would be to present the stimulus only to him. Unfortunately, this is just what cannot be done in parapsychology, since there is currently no way known of screening anyone from being affected by, say, a given pack of ESP cards.

From the very earliest days of research it has been clear that the very existence of, for example, telepathy, would make it impossible to know for certain whether the result in a psi experiment was due to the mechanics of the effect itself, or whether the subject merely found out what the experimenter wanted to happen by telepathy and acted to confirm his desire. An anarchic phenomenon such as psi, with no known limitations, would by its own nature render improbable any exposure of its inner secrets.

The observational theories have taken seriously the empirical suggestions that psi is independent of space and time and thus act to dampen any hope still springing in the parapsychologist's breast, that these variables can be used for the purpose of screening off unwanted influence. At the same time these theories have as yet yielded no other certain physical limitation which can be used to curb the unruly nature of psi. They do, however, suggest the possibility that ultimately suitable feedback manipulation may provide our Alladin's lamp. For the moment, though, they have made the darkness denser by replacing the ungainly and improbable early conceptualizations of experimenter influence by an elegant scheme in which the experimenter is merely another subject affecting a random generator. The RNG in this case is the whole experiment. No longer do we make laborious schemes "S gets E's expectations by telepathy and reacts so as to honour them" but instead "Psi is goal directed: if E has psi and he wants the experiment to produce a particular result, then he will use his PK to bring that about".

We are therefore forced to use other, relatively uncertain, means for deciding who the subject is. The primary one of these is that psi effects appear in association with certain individuals. Criteria secondary to "psi source by association" can be formulated, such as that effects are strongly correlated with the psychological state of one person. The point which should be emphasised here is

that, at present, only on these weak correlational grounds can we say that psi effect X is produced by person P.

If psi effects appear around one person we may suspect that he is a psychic. On this basis it seems fairly certain that the effects produced in association with high-scoring subjects, such as Stepanek (Pratt, 1973) are exactly what they seem. His results are (virtually) independent of the identity of the experimenter but on the contrary, are sufficiently associated with Stepanek that a number of parapsychologists have flown half way around the world for the privilege of working with him. It seems likely that it is a characteristic of high-scoring subjects in general that the results are singularly associated with them and rather independent of the experimenter.

When we examine the typical successful group experiment using unselected subjects, however, a totally different picture emerges. Scoring is not associated with individual subjects: indeed such experiments typically have no individual high scorers, though the result, as a whole, is significant. On the other hand, results are conspicuously associated with particular experimenters. In the last few years Honorton, Schmidt, Braud and Stanford have been among the most prolific of these.

On the basis of the major criterion we have for judging ownership of a psi effect, there is the suggestion that these special experimenters are actually psychics in disguise. However we conceptualise the fact of experimenter differences, though, it has important implications for the distribution of psi. Most workers using group experiments either find no evidence of psi at all, or alternatively marginal effects which do not recur on replication. Experimenters of this type (e.g. Pratt, 1977, p 75) do, however, sometimes find high-scoring individuals. In strong contrast, the results of the special experimenters, using groups, give statistically strong evidence of democratic psi which is fairly robust on replication: they too find occasional special subjects. In other words the literature suggests that special experimenters find the Schmeidler-McConnell distribution, while the rest observe only the psi-star component. It is clear that any empirical attempt to determine psi distribution must take account, and rather powerful account, of the identity of the experimenter.

It is of the utmost importance to determine whether the difference between experimenters is due to experimenter-psi or the normal sensory-mediated experimenter effect: we know in principle how to control for sensory-mediated effects: if experimenter psi is involved, on the other hand, parapsychology is in deep trouble (Broughton, 1979).

The first thing to determine is whether there really are only a few special experimenters: those who are in the minority should be most carefully inspected. McConnell (1977, p 203) has commented how dependent parapsychology is on a very few experimenters " .... in the years 1970-1975 more than 50% of all the statistically evaluated experimental papers in the Journal of the American Society for Psychical Research were authored or co-authored by one of five parapsychologists." Rhine (ATV, 1976) has given his opinion on the frequency of experimenters who can obtain evidence of psi "... we have found that very few people can succeed as investigators, very few people can go out and get it ....". There seems to be a concensus that a handful of special experimenters bring home most of the parapsychological gravy.

It has been known for a long time that experimenters working with the same subjects get different results (White, 1976; Kennedy and Taddonio, 1976) but this has usually been attributed to the psychological skills of the experimenter in coaxing psi from subjects. Rhine (ATV, 1976) continues ".... no matter how many degrees they have, how much education, how smart they are in many ways, there is a kind of smartness, there is a kind of intelligence, a social intelligence that they've got to have before they can get people to perform right ....". Pratt joined with Rhine in an extended treatment of the importance of social interactions (Rhine and Pratt, 1976, p 131-138). It is rather chastening to find that one so capable of these human insights should yet himself fail so dismally to elicit psi in unselected subjects (Pratt, 1977, p 75).

The hypothesis that experimenter differentials are due to psychological variables has been widely accepted by parapsychologists. In the early days there was, of course, no alternative explanation available: it just had to be psychological. However, there is surprisingly little hard evidence. A detailed discussion of the rival merits of the sensory-mediated versus the psi-based models is outside the scope of this paper. However a brief résumé may be in order.
(1) Parker (1977) found that successful experimenters scored in a psi task disguised as a personality inventory, while the unsuccessfuls did not. Sargent (1979b, p 177) also claims confirmatory unpublished data. This does not, however, absolutely distinguish the two hypotheses since it is conceivable that an experimenter psychologically able to elicit psi in others may do so for himself, perhaps even when he is not trying to.
(2) Parker (1977) also looked for personality differences with the 16 PF, but failed to find any. Sargent (1979a), restricting himself to ganzfeld experimenters, did report differences indicating that successfuls are significantly more sociable (A), extravert (QI) and enthusiastic (F) but less neurotic. Parker (in press) has, however, described Sargent's work as "little more than a dressed up case

study in which results are obtained by fitting the data to the hypotheses.".

Even if personality differences were to be established this, per se, would have little bearing on whether the experimenter effect is sensory or psi-based. Personality differences might be relevant only in so far as they were correlated with experimenter psi differences. It would be necessary to employ multivariate (here bivariate) methods to decide which came first.

However, an explanation of the extraordinary success of our few special experimenters in terms of the kind of psychological variables measured, seems somewhat implausible. Unless the personality characteristics associated with them were similarly extremely rare, they could not possibly account for a large part of the variance.

(3) Experimenter differentials are obtained even when subjects and experimenter never meet (e.g. West and Fisk, 1953). Since contact is necessary for a sensory-mediated effect, these results at least, cannot be due to a sensory influence.

(4) If E-effects are due to social interaction then one might expect experimenters to become more skilled at eliciting psi as time goes on. In fact special experimenters, like special subjects, generally decline. Schmidt, for example, cannot now "elicit" psi from his subjects. It might be counter-argued that experimenters are likely to become progressively less enthusiastic and thus counter any learning incline. It seems difficult, however, to reconcile this with the long-term failure of centres such as the FRNM to train enthusiastic youngsters to elicit psi.

(5) The data of different subjects collected by the same experimenter tend to show the same "signs of psi", for example particular displacements (West, 1953; Sargent, 1978). It appears somewhat strained to suggest that E subtly cued subjects to make, say, -1 displacements.

(6) A few workers have tried directly to manipulate treatment of the subjects. Perhaps the best-known of these is Honorton, Ramsey and Cabibbo (1975). Briefly, one group of subjects was treated nicely and the other badly. The nicely treated group scored significantly above chance, as expected, and there was a significant difference between the scoring rates of the two groups. However, the badly treated group scored significantly below chance to about the same degree as the positively - toned group scored above.

If we take this result to heart the best thing we could do would perhaps be to treat subjects as abominably as possible, so that significant, if negative, scoring could be secured. Further, the effect was restricted to only one of the two experimenters. Thus, while the result was evidently due to an experimenter influence, the effect of the subject treatment, was rather less important than the

identity of the experimenter.

The experiment, moreover, made no attempt to determine whether the result was due to sensory interaction or E-psi. To be sure, at present there is no certain way to do this, but at least a start could be made. A simple addition would have been a condition in which there was no sensory contact between S and E. In general, these studies treat the problem of E-psi or E-psychology by the debatable expedient of ignoring the psi possibility.

PROPOSED SCHEME FOR PSI DISTRIBUTION

On empirical grounds the view of sensory-mediated experimenter influence has no striking advantage. On the contrary, the experimenter-psi idea seems to be a nose ahead. The sensory picture does, however, have one advantage: it seems so much more intuitively likely. On the other hand the E-psi view has hardly ever been discussed and is relatively unfamiliar. The bulk of this section will therefore be devoted to outlining a scheme of psi-distribution involving E-psi and developing a few of its ramifications to the point at which testable consequences can be drawn.

The scheme proposed is that there is no generally distributed low level psi ability in the population: the real distribution is a psi-star one. The few experimenters who themselves have psi use their subjects simply as random generators: their process-oriented studies are therefore likely to reflect more about their own hopes and fears than to reveal anything about psi. The remaining experimenters see only the psi-star component.

A host of questions spring to mind.
If it is true that in successful experiments the experimenter is actually contributing the psi might one not expect all his subjects to score at about the same level?
It has often been noted that this is actually observed. Schmeidler and McConnell (1958, p 96), for example, comment "There is little inhomogeneity within or among our subjects.". It is current practice to perform single-mean t-tests (Stanford and Palmer, 1972) to ensure that results are homogeneous over subjects and they almost always are. If the scoring rate were higher this would be a strong indication that something other than subjects was causing the results: as it is, this finding has been tacitly attributed to the low power of the tests to detect differences round about the typically low scoring level. However, some differences are to be expected between subjects in ESP tests in their responsiveness to experimenter PK. If we invoke the brain RNG picture discussed earlier the subject differences would be partly independent of whether the subject himself, or the experimenter, was providing the

PK input. It could be expected that for experimenters with psi, any manipulation of subjects which made them better random generators would result in improved scoring. For an experimenter without psi, whether his subjects were good RNGs or not would be irrelevant, since there is no PK influence to work upon them. In other words appropriate social interaction might be of the essence for a special experimenter, yet be of little or no importance to experimenters without psi.

Even if the experimenter is the psi source, why should one be suspicious of the answers he obtains in process-oriented experiments? After all, it is genuine psi regardless of its origin. Maher and Schmeidler (1978), p 391-392) have put this with some eloquence "Much concern among parapsychologists revolves around the issue of psychokinetic 'experimenter effect'. However, ... experimenter PK presumably can also be used in an effort to derive nonegotistical answers to legitimate theoretical questions ....". This is possible but seems somewhat implausible in that it ascribes moral rectitude to actions which are not under direct conscious control. For experimenters with psi their dream world becomes objectified, by PK, to a very tiny extent, in external reality. It is then, as if we were to put our faith in the validity of an experiment done in a dream. Palladino was, presumably a psi-source too.

If the experimenter is the origin of the psi, why does he not always get the results he wants? In practice our special experimenters do usually obtain results rather close to their expectations (e.g. Honorton's muscle tension experiments, 1976, p 219). That they do not always succeed entirely, but obtain something related, or indeed the reverse, is no more remarkable than that other special subjects should sometimes make displacements or psi-miss.

Why should there be a relatively higher concentration of special subjects masquerading as experimenters than there are special subjects in the general population? The proportion of special experimenters is actually a rather small fraction of the number of serious parapsychologists: there are, at the outside, ten special experimenters in the Parapsychological Association, which has something like 200 members and associates. On an absolute basis, then, that is 5%. This is still probably many times the frequency of psi-star subjects in the general population. It would be surprising indeed if professional parapsychologists were not a very special self-selected breed. One of the factors involved in the process is clearly the ability to get results. Rhine and Pratt (1967, p 132) make no bones about it "A psi experimenter is one who, under conditions that insure he is not fooling himself, can get results. All others should do something they can do well.".

If the special subjects who are our special experimenters obtain scoring from their subjects, then should it not be the case that undisguised special subjects are able to elicit psi from an unselected group?

There are many claims that special subjects are able to transfer

their powers, for example with Bill Delmore (Kelly and Kanthamani, 1973), but there has been little experimental investigation. A recent study (Palmer and Tart, 1979) looked for psi transfer with Matthew Manning, but failed to find any convincing evidence of it. The authors, however, present no evidence that Manning was able, at that time, to produce the psi effects which were required from the hopeful recipient of his powers. One of Rhine's early talented subjects, Stuart, did produce significant scores using unselected people: further, it is interesting to discover that while Pratt was himself scoring successfully as a subject, he also "elicited" scoring from students (Rhine, 1964). However, the experimental conditions employed, would not pass muster by current standards.
Is it reasonable to assume that experimenter-psi remains about constant?
On the contrary subject psi sometimes erupts suddenly in someone for a few weeks and then vanishes forever (e.g. Mitchell, 1953). Similarly there are suggestions that a number of experimenters have success in their first experiments and then go into rapid decline. It is not improbable, however, that this is a reporting artifact. The psi ability of special experimenters seems to last for longer than that of special subjects. One might speculate on possible psychological reasons for this. Schmidt is now well into decline and it might be argued that Stanford's results are no longer what they were. Braud is at his peak. In terms of longevity, though, on this picture of psi distribution, Honorton must surely be the Stepanek of the experimenter set, and he still shows no signs of decline.

In addition to the long-term comings and goings of psi ability, there are short term variations. At the 1976 Parapsychological Association convention in Utrecht Schmidt talked about his "dry times" when he could "elicit" psi from no-one, presumably including himself.
If the psi we see in groups comes from experimenters but that of special subjects comes from themselves, then would one not expect to see evidence of two kinds of psi?
In terms of personality research Schmeidler (1974, p 101) has put it beautifully "We need two strategies: full personality studies of the best, authenticated psychics to see what commonality they show and comparable studies of others for whom we have ESP or PK scores. If the two sets of findings mesh, we accept the second possibility above, a continuous range in psi ability. If they are discrepant, we accept the first (a qualitative difference).". Much more research is needed but I cannot help quoting Koestler (1974, p 167) ".... ironically, Rao's composite portrait of the potentially high scoring ESP subject .... is in almost every respect the exact opposite of Pratt's description .... of the highest scoring person known at present - Pavel Stepanek.". If the ESP process is indeed a two stage one, as suggested, then perhaps the characteristics

associated with having good connection to one's psi source are quite different from those associated with having a good brain RNG. At least this suggests some quite exciting psychological prospects for investigation.

The proposal has been written in the form of one particular person having the psi: is it not more reasonable to suggest that people combine, as Kennedy and Taddonio (1976, p 22-23) imply? This seems highly probable and somehow, it sounds much less radical, but in fact it is only a rather minor modification. If we once conceive of the experimenter having a "psychic finger in the pie", it ceases to matter much how tiny a finger it is: he is still going to pull out the plum. More explicitly, this modification does not increase our confidence that our special experimenters' process-oriented results are meaningful. Secondly, to produce interpretable data we must either learn to detect where the psi is coming from, or be able to screen the experimenter portion off, just as with the unmodified proposal.

Will the psi-star component be the same regardless of the psi ability of the experimenter looking? The answer to this is dubious. In the absence of a theory of psi nothing can be said. If we wish to adopt the observational theory view then the answer is probably, no.

According to the observational theories, the function of PK is to change probabilities, in other words psi is an alternative name for luck, the kind of luck that sticks with a given person. Consequently, an experimenter will not only be lucky in having his experiment turn out according to his hopes (using his own PK) but will also be lucky in finding the kind of subjects he needs. In other words, an experimenter with psi will find (if he wants that) many more subjects of the psi-star type than would an experimenter without psi. At first sight this seems more hopeful for process-related investigation than a group experiment with a special experimenter; but this may not be true. If the experimenter's desire is to show, say, that standing on the head improves psi, then his goal-directed PK will act so as to satisfy that need, selecting only special subjects who will show just that behaviour. There is some empirical justification for the suggestion that special experimenters find more psi-star subjects, for example compare Fisk to West (West, 1973; 1954).

Is there any real evidence that special experimenters have psi? If the scheme proposed above is correct then experimenters with psi will get their experiments to work, while those experimenters with no psi ability will obtain only chance results. The micro-PK review by Honorton (1978) provides a convenient base on which to test the prediction. Most of the experimenters involved have considerable experience in the field, thus data about them are available.

How can an experimenter's psi ability be determined from the literature. The only practicable method is to scan for work in which

these researchers have scored significantly as subjects. It is probable that this is a rather poor criterion for psi ability. It will be heavily contaminated, particularly by reporting bias: it has generally been regarded as somewhat improper for an experimenter himself to have psi: however the probable result of this will be to attenuate any relationship which exists, rather than to inflate it. If any of the experimenters in a multi-author paper qualify as psychic then it is classed as "psychic experimenter". For literature references see Millar (1978). (Note: in that paper one experiment was incorrectly classified.) An experiment is classified as successful if the result is significant at or better than the 5% level. The data are presented as Table 1.

TABLE 1

Relationship between psi ability of E and the success of his micro-PK experiments

	Success	Failure	Total
Psychic E	33	9	42
Residue	2	10	12
Total	35	19	54

Chi-square (corr) = 13.1 df=1 $p<3x10^{-4}$

Psychic experimenters have significantly ($p<3x10^{-4}$) more experiments succeed (79%) than the others (17%), as predicted. It is, moreover, rather surprising to find that psychic experimenters account for over 3/4 of all the studies. The scoring rate for the residual (non-psychic) experimenters is likely to be considerably inflated by the usual bias towards publishing only positive studies. Honorton (1978, p 9) takes an outside estimate that 10 insignificant and unreported studies occur for every one reported. More plausible would be a ratio of 2 to 1. In this case, for the non-psychic group, the two successful studies would actually be selected from 36, rather than the 12 reported: this is 5.6%, whereas chance expectancy is 5%. It is unfortunate that the last step involves a guess, however reasonable it may be. It is interesting, therefore, to calculate the probability of the observed number of successful experiments for the non-psychic group occurring by chance alone. The exact binomial probability of two or more such

studies from the reported 12 tries is 0.12, not significantly different from chance expectancy. The data strongly suggest, that at least in micro-PK, if an experimenter does not have publishable psi himself, then he is not likely to see any in his experiments.

In summary, a scheme has been proposed relevant to the distribution of psi in the population, one which contrasts rather strongly with the usual view. In this scheme, only a few have psi ability and of these some are parapsychology's special experimenters. The picture is coherent and in accord with much experimental evidence. Further, it seems to have some predictive value. Only experiment can, in the last resort, determine how much or how little truth it contains.

EMPIRICAL DISTRIBUTIONS

Schmeidler and McConnell (1958) present the distributions of total scores for subjects each performing 9 runs with ESP cards. These are reported separately for sheep (889) and goats (363), the data have been summed to reconstitute the parent population (889 subjects), giving Table 2 and figure 5.

TABLE 2

Observed and expected distributions of Schmeidler and McConnell's subject-total scores, for the 889 subjects (526 sheep, 363 goats), each performing 9 runs

run total	0-34	35,36	37,38	39,40	41,42	43,44	45,46
number observed	27	46	62	79	98	119	96
number expected	32.6	34.5	56.0	80.3	102.1	115.4	116.7
chi-square (1 df)		3.99x					4.23x

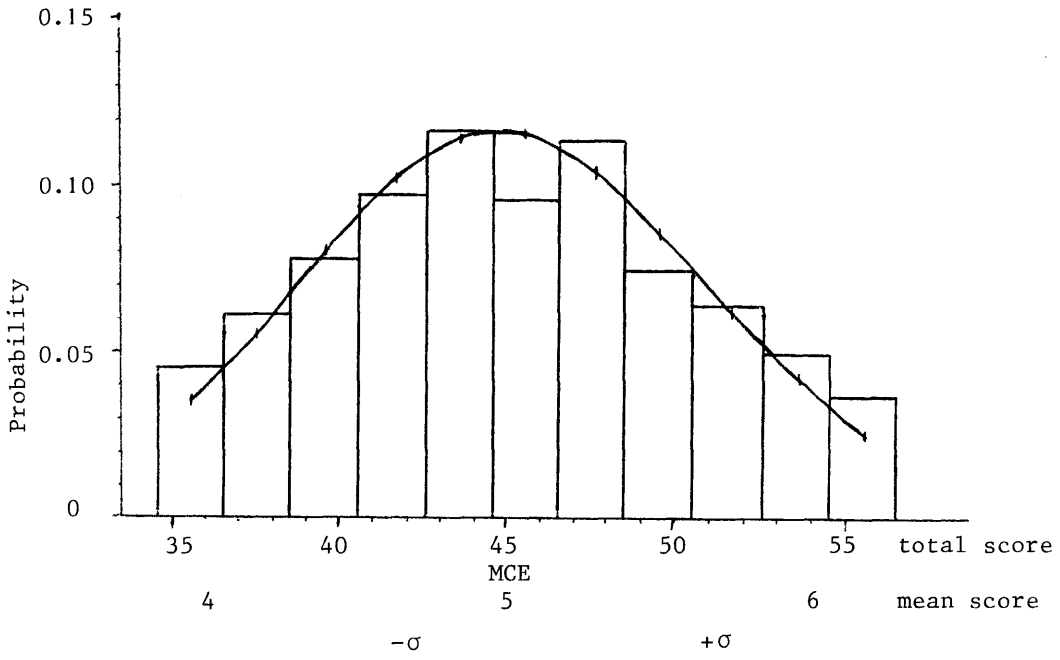
	47,48	49,50	51,52	53,54	55,56	57 and over
	115	77	65	50	37	28
	105.9	86.5	63.7	42.6	25.8	26.9
					5.01x	

x $p<0.05$

chi-square = 17.5 df = 12 ns

Overall, the curve conforms to binomial expectation. However, 3 of the 11 histogram bins are independently significant. The two extremes

Figure 5. Scoring distribution for sheep plus goats



are over-represented. On the low-score side this is due to the goats, while on the high it is the result of the sheep. Furthermore chance scores are under-represented. This is due almost entirely to the goats.

The last would hardly be worth mention save that a similar effect has been noted elsewhere, but in a run-score distribution with a single subject. Bierman (1978b) observed under-population of chance scores in computerised precognition self-tests, in both of the conditions used, as well as in old data. Bierman comments that such an effect is to be expected if the subject has a fixed amount of information which he may add to the system and he employs it in targetting for run-score instead of for hits on a trial-by-trial basis.

A micro-PK experiment (Broughton, Millar, Beloff and Wilson, 1977 and in preparation) used pre-recorded targets on cassette tape. The results for all experimenters were homogeneous and the data for the 256 subjects were pooled. However, the experiment gave no evidence of psi and thus, not surprisingly the distribution conforms to binomial expectation.

Walker (1975) enterprisingly assembled a distribution for 1002 subjects from 14 different experiments reported in early volumes of the Journal of Parapsychology. Unfortunately, he simply assumes that these studies can be meaningfully pooled. There is no attempt to determine whether the experiments have homogeneous distributions: this is quite improbable since they varied in such important factors as experimenter, type of test, number of trials and p-hit. He attempts to correct for the p-hit differences by using, not the scoring rate, but the genuine hit probability; however, this is a shot in the dark. Furthermore Walker's plot is "corrected for those expected in the original data from chance alone". To subtract the binomial expectation for those subjects who scored only at chance requires their number. Walker does not explain how this estimate was made.

At this stage meaningful pooling of disparate experiments is not in general possible: the unit of analysis, for the present, must be the individual experiment (or indeed only the part of an experiment which can be considered homogeneous). In order to obtain a distribution of subject scores with some statistical stability it is necessary (a) for each subject to perform a reasonable number of trials and (b) a large number of subjects should be tested. The number of experiments in which, say, more than 100 subjects are tested under standard conditions is rather small, thus partly accounting for the rarity of empirical distributions in the literature. While there is still little knowledge of how psi may affect the distribution

these large-scale experiments are primarlily important in providing suggestions to visual inspection, of the kinds of thing that may be going on.

However, Walker's approach of using much smaller experiments to find out about distributions, while premature, holds out more promise in the long term. If parapsychological workers routinely reported subject distribution data (as was done much more frequently in the early days) then, even though each one was extremely contaminated by noise, tests could be made on the set of curves. For example, suppose we take the suggestion from Schmeidler's data that chance-scores are under-represented. If we have 20 curves from small-scale experiments the Z-scores of the middle bins can be compared with a single mean *t*-test against the expected mean of zero.

In addition to the above large-scale curves reported, two recent experiments have collected sufficient data to yield distributions, though in neither case has this actually been done as yet. The first is a large-scale classroom screening exercise reported by Palmer, Tart and Redington (1976). This employed about 2000 subjects in all. There were 10 cells in the experimental design, thus giving about 100 subjects per cell. While there were differences between cells the overall mean and variance did not differ from chance expectation.

The chance values of the first two moments are consistent with the possibility that if there are enough conditions, subjects, experimenters in an experiment, the overall distribution curve will approach chance expectation. The few psi-star subjects out on the tails will be diluted into overall insignificance by the residuum. This would be consistent with the large-scale trials carried out each day by the gambling industry. It is, furthermore, entirely compatible with Braud's "spreading thin" idea (Wood, Kirk and Braud, 1977; Braud and Braud, 1979), which in its turn follows immediately from Walker's theory (analogous to run-score targeting: see Millar, 1978), if it is assumed that the experimenter is the major source of psi in his experiments.

The other set of data which have not yet been published in full is the Amsterdam fast RNG experiments of Bierman and Houtkooper (1978). A standardized micro-PK task was attached to each of ten different experiments: in total 92 subjects were employed. The strategy of combining standard screening trials with on-going experiments might be considered by other experimenters interested in populations.

Thus of the only two curves currently available to the writer, one of them (Schmeidler and McConnell) shows only the democratic

psi distribution (bi-directional here) and the other (Broughton et al.) shows no psi at all.

In addition to distribution curves it is of considerable interest to determine the incidence of psi-star subjects. A thorough survey of the literature, while desirable, is outside the scope of this paper. However, some estimate of the frequency should be made here. Some selected examples of screening experiments are Fisk, 1951 and follow-up Mitchell, 1953; West, 1947; Randall, 1974a and follow-up, 1974b; Tart, Palmer and Redington, 1979. Estimates of the prevalence of psi ability range all the way from 6% to 0%. The last study may perhaps be selected on account of the large number of subjects employed. From an initial batch of 1835 subjects, 3 out of the 10 finalists scored significantly in the training study, 2 at the 5% level and one at 0.4%. Because of the large number of drop-outs the 0.16% of psychics may be conservative. From comparison Schmeidler and McConnell (1958) found no consistently high-scoring subjects among their 1308 subject sample. As a very rough estimate perhaps one in a thousand may be a psi-star subject.

DISCUSSION AND RECOMMENDATIONS

The above has dealt with psi distribution in a narrow sense; however, the ideas outlined, particularly with respect to parapsychology's special experimenters, have more far-reaching consequences.

The most objectionable feature of the experimenter-psi scheme is that almost the entire literature of the past 50 years must be discarded. Some possibilities might be suggested to retain the good features and discard the unpalatable ones. For example, could it not be that special experimenters are quite ordinary and that those who do not see evidence of psi are actively inhibitory? This does not seem likely: high-scoring subjects are not often extinguished by even the most psi-null experimenters (e.g. West and Fisk, 1953). In any case there is direct evidence that the psi-plus group has psi and in terms of frequency they are certainly the odd men out.

What about the possibility that psi-plus experimenters simply release their subjects' psi, presumably via the experimenter's PK. At first sight this seems an excellent compromise. However, when we note that at least in some cases (e.g. Honorton, 1976) experimenters almost certainly used their own psi to produce the desired result, the confidence this proposal suggests immediately collapses. If experimenters have psi at all, their results cannot be taken at face value.

An unfortunate result of the importance of the experimenter, for whatever reason, is that experimenters are well aware that their

results are critically dependent on their own mood and enthusiasm. When most special experimenters are not in a fired-up state no scoring is obtained. Consequently, experiments are hardly ever repeated exactly and the hum-drum repetitive jobs that are the daily bread of most psychological work never get carried out. Even if the experimenter effect is a conventional sensory one, if experimenters are always forced to jump from topic to topic, no really solid data-base can ever be laid.

If an experimenter effect, of whatever kind, is the over-riding factor in psi experiments then something has to be done. In short, now that we know we have it we must get rid of it. Any experiment with only a single experimenter is not worth very much, because of the possible sensory contribution. If the experimenter effect is a sensory one, however, it is relatively easy to minimise it: a number of data collectors should be employed. At the very least, a routine feature in psi experiments should be that someone else runs some of the subjects. However, this kind of precaution is of little use against experimenter-psi. Sometimes lack of sensory contact with the principal investigator seems to block his contribution (e.g. Pratt and Price, 1938) but more often it does not (e.g. Fisk, 1951). Perhaps this manipulation has an effect on the experimenter's psi contribution only if it changes his attitude.

Can the observational theories be of any help in detecting or eliminating experimenter-psi? At the present state of the theories no cut and dried answers are possible. It would be surprising if they could provide any certain prescription for parapsychology's ills without the aid of experiment. As in other scientific work the theory should suggest experiments which lead to a more useful theory and so on. Some effort has been expended on the theories in the hope of improving on them in this respect (e.g. Millar and Hartwell, 1979) and this work continues; however in the virtual absence of empirical evidence this is very much a guessing game.

However, the theories do offer strong suggestions. In particular they assert that feedback is a necessary condition for anyone to exert a psi influence. Consequently, if the feedback and/or the experimenter, is manipulated at feedback, then his contribution should change. One complication is that E typically obtains feedback at a large number of points in the experiment. Thus, while it is likely that he usually exerts his major PK effect somewhere near the final analysis, for the purpose of experimental feedback manipulations it would be advisable to limit his contact with the data so far as is practicable, so that there is only one point of feedback. It will be noted that this is consistent with the parapsychological lab lore that the experimenter should not examine his data until they are all in and then only when he is in a good mood.

The modification suggested by the observational theories is that the data should be split, unseen, into two subsets. For example E then looks at one set when he is in a good mood, the other when in a bad mood: he could be drunk or sober, in ganzfeld or not etc. If experimenter psi is an important factor then these manipulations will probably affect the results obtained. It would be advisable to adopt this general procedure as an integral part of all psi experiments. The easiest process would be that two different people observe the data.

Such manipulations do indeed produce scoring differences (e.g. Broughton, 1977; Bierman, 1978a). What exactly these differences mean, however, and how we can proceed further, are not so clear. The experimenter (or someone else) may look later at the difference between subsets, which he eventually does see. Because the observational theories are formulated so that psi is space-time independent he may possibly use his PK then and treat the whole thing just like any other experiment. It is clear, however, that factors in the real world limit this space-time independence and the task of experiment is to determine what these may be. If these factors can be isolated then they can be deliberately used to control experimenter psi.

What sort of evidence would suggest that a subject was playing some part? A strong indication would be if the subject could score with other experimenters. If the experimenter himself were unable to score but could do so only in combination with a particular subject, this would strongly suggest the importance of the subject. In general if different subjects scored at grossly different rates with the same experimenter, this would suggest the subjects had some influence, especially if the scoring rates were much in excess of those the experimenter could produce unaided.

The question of the combination of psi sources is of considerable importance (Millar, 1978, p 323). A most interesting case would be if neither S nor E could score separately, yet they could when acting together. If such an effect should prove to be other than a psychological effect, then the suggestion would be that a psi source was being created from two non-sources.

It might be thought that the ideas proposed here would be alien to our special experimenters. On the contrary, these workers seem always, like poltergeist agents, to have harboured some deep suspicion about their own involvement in the results they observe. However, it is only lately (e.g. Honorton, 1975) that this has attracted such attention. There are a number of comments in the literature which come very close to the position outlined in this paper. Having accidentally noticed that if data are checked by

different people the results are different (i.e. what the observational theories suggest) Feather and Brier (1968) comment "If this situation is true, the checker in the present may have an effect upon the calls which the subject made in the past". It can only have been due to the weight of pre-conceived opinion that they did not think to ask "If the checker is more important to the results than the 'subjects', whom should we be studying?".

One of the basic assumptions in parapsychology has been that psi is dependent on a constellation of delicate psychological factors. After 50 years of attempts to define these variables, parapsychology has little or no more control over the phenomena (Pratt, 1978, p 135-137). It is the contention of this paper that the failure of parapsychology to progress, results rather from a fundamental misconception about the psi process, in particular where the psi is coming from. Whether this is true, or not, will ultimately be determined by whether the ideas expressed here lead to greater control over psi effects.

Most experimenters are handicapped by the inability to produce any psi to study: on the other hand what they see probably reflects the true state of play in the world. One possibility for this majority is to initiate large-scale screening projects to detect the psi-star subjects. Another is, in effect, to use special experimenters as subjects. A compromise might be to obtain the co-operation of a special experimenter to help with screening, since, as noted above, his psi should attract others.

For the special experimenter, it would be premature to discard his current theories for untried ones. Diversity of view, rather than unity is more likely to prove productive in parapsychology. Nonetheless, it seems the special experimenter would be well advised to "take out insurance" against his own decline by some investment in feedback manipulation.

ABSTRACT

The topic of this paper is the distribution of psi in the population. Although the basic strategy of experimental work depends upon this, surprisingly little attempt has been made to examine the question experimentally. Instead, since the Rhine revolution, it has been assumed that most people have some small degree of psi ability. The assumption is questioned here.

Drawing heavily upon the observational theories, and Walker in particular, it is suggested that the principal psi differences between people are determined at feedback. The degree to which the feedback stimulates the psi source is proposed to be crucial. The

main experimental consequence is that suitable manipulation of the state of the subject at feedback may lead to improve psi scoring.

If psi exists, the traditional distinction between subject and experimenter can no longer be drawn. The results of special subjects are almost independent of the experimenter, while in group experiments he is crucial: only a handful of special experimenters can elicit psi from unselected groups with any regularity. Explanation of this in terms of experimenter psi is decidedly no worse than the usual idea of sensory-mediated effects. Experimenter psi has also the logical advantage that it alone can be tested (by removing the possibility of a sensory component). In a few cases this procedure has left experimenter differentials intact. This section is independent of the observational theories.

The problem with experimenter psi is not any logical or empirical deficiency, but rather what to do with it. To this end a scheme of psi distribution is proposed which uses the observational theories to suggest ways out of parapsychology's dilemma. The reason most experimenters do not obtain convincing evidence of psi with unselected groups, is that there is none. Instead, this is an artifact of the psi abilities of the few special experimenters. This scheme accounts for many of the observed features of parapsychological experiments and some novel deductions can be made. Only psychic experimenters should be able to get group experiments to work: this is strongly supported by the micro-PK literature.

On this basis, the majority of experimenters who cannot "elicit" psi in the conventional way, will never be able to and should turn their attention to attempts to find special subjects or use special experimenters as subjects. Special experimenters, on the other hand, should attempt to manipulate their own psi contribution by working with the feedback they get, within the context of the observational theories.

ACKNOWLEDGEMENTS

I owe considerable intellectual debt to Harris Walker and Helmut Schmidt: they are, of course, entirely blameless for the use I have made of their work. Correspondence with Gertrude Schmeidler and Jurgen Keil was invaluable in sparking off this paper. I am grateful to Lawrence Moore for the transcript of "Into the Unknown". My colleagues at the Utrecht parapsychology laboratory have been a constant source of inspiration.

REFERENCES

ATV Network Ltd. Excerpt from transcript of "Into the Unknown", 1976, producer Lawrence Moore.

Bierman, D.J. Observer or experimenter effect? A fake replication. E.J.P., 1978, 2, 115-125 (a).

Bierman, D.J. Testing the 'advanced wave' hypothesis: an attempted replication. E.J.P., 1978, 2, 206-212 (b).

Bierman, D.J. & Houtkooper, J.M. The quest for reproducibility: a survey of Amsterdam fast RNG experiments. J. Parapsychology, 1978, 42, 56-57, abstract of SERPA paper.

Braud, L. & Braud, W. Psychokinetic effects upon a random event generator under conditions of limited feedback to volunteers and experimenter. J.S.P.R., 1979, 50, 21-32.

Broughton, R.S. An exploratory study on psi-based subject and experimenter expectancy effects. R.I.P., 1976, 1977, 173-177.

Broughton, R.S. Repeatability and experimenter effect: are subjects really necessary? Parapsychology Review, 1979, 10, 11-15.

Broughton, R.S., Millar, B., Beloff, J & Wilson, K. A PK investigation of the experimenter effect and its psi-based component. R.I.P., 1977, 1978, 41-48.

Dorfman, D.D. The Cyril Burt question: new findings. Science, 1978, 201, 1177-1186.

Feather, S.R. & Brier, R. The possible effect of the checker in precognition tests. J. Parapsychology, 1968, 32, 167-175.

Fisk, G.W. Home testing ESP experiments. J.S.P.R., 1951, 36, 369-370 and 518-520.

Honorton, C. Has science developed the competence to confront claims of the paranormal? Presidential address to the P.A., R.I.P., 1975, 199-223.

Honorton, C. Replicability, experimenter influence, and parapsychology: an empirical context for the study of mind. Presented at the annual meeting of the A.A.A.S., Washington, 1978.

Honorton, C., Ramsey, M. & Cabibbo, C. Experimenter effects in extrasensory perception. J.A.S.P.R., 1975, 69, 135-149.

Kelly, E.F. & Kanthamani, B.K. A note on a high-scoring subject. R.I.P., 1972, 1973, 84-86.

Kennedy, J.E. & Taddonio, J.L. Experimenter effects in parapsychological research. J. Parapsychology, 1976, 40, 1-33.

Koestler, A. Postscript pp 165-169 in New directions in parapsychology. John Beloff (Ed.), Elek, London, 1974.

Maher, M. & Schmeidler, G.R. Letter to the editor. J.A.S.P.R., 1978, 72, 389-392.

McConnell, R.A. The resolution of conflicting beliefs about the ESP evidence. J. Parapsychology, 1977, 41, 198-214.

McNicol, D. A primer of signal detection theory. George Allen and Unwin, London, 1972.

Millar, B. The observational theories: a primer. E.J.P., 1978, 2, 304-332.

Millar, B. & Hartwell, J. Dealing with divergence. R.I.P., 1978, 1979, 91-93.

Mitchell, A.M.J. Home testing ESP experiments: special report on one series of tests. J.S.P.R., 1953, 37, 155-164.

Palmer, J. & Tart, C.T. Delayed PK with Matthew Manning: preliminary indications and failure to confirm. E.J.P., 1976, 2, 396-407.

Palmer, J., Tart, C.T. & Redington, D. A large-sample classroom ESP card-guessing experiment. E.J.P., 1976, 2, 40-56.

Parker, A. Parapsychologists' personality and psi in relation to the experimenter effect. R.I.P., 1977, 107-109.

Parker, A. In defence of introverts: a critical note on supposed personality differences between ganzfeld experiments. E.J.P. (in press).

Pratt, J.G. A decade of research with a selected ESP subject: an overview and reappraisal of the work with Pavel Stepanek. Proc. A.S.P.R., 1973, 30.

Pratt, J.G. Parapsychology: an insider's view of ESP. Scarecrow Press, Metuchen, N.J., 1977.

Pratt, J.G. Prologue to a debate: some assumptions relevant to research in parapsychology. J.A.S.P.R., 1978, 72, 127-139.

Pratt, J.G. & Price, M.M. The experimenter-subject relationship in tests for ESP. J. Parapsychology, 1938, 2, 84-94.

Randall, J.L. Card-guessing experiments with schoolboys. J.S.P.R., 1974, 47, 421-432 (a).

Randall, J.L. An extended series of ESP and PK tests with three English schoolboys. J.S.P.R., 1974, 47, 485-494 (b).

Rhine, J.B. Extra-sensory perception. Bruce Humphries, Boston, 1964.

Rhine, J.B. The reach of the mind. William Sloane, N.Y., 1971.

Rhine, J.B. & Pratt, J.G. Parapsychology: frontier science of the mind. Charles C. Thomas, Springfield, Ill, 1967.

Sargent, C.L. Experimenter psi-effects II. A possible 'experimenter mindprint'. E.J.P., 1978, 2 239-246.

Sargent, C.L. Repeatable significance and the significance of repeatability. Paper read to the Third International Conference of the S.P.R., Edinburgh, 1979 (a).

Sargent, C.L. The Parsons experiment with Basil Shackleton: some neglected data. J.S.P.R., 1979, 174-179 (b).

Schmeidler, G.R. The psychic personality. In Psychic exploration a challenge to science, John White (Ed.), Putnams, N.Y., 1974, 94-110.

Schmeidler, G.R. & McConnell, R.A. ESP and personality patterns. Yale University Press, New Haven, 1958.

Schmidt, H. Towards a mathematical theory of psi. J.A.S.P.R., 1975, 69, 301-319.

Schmidt, H. & Pantas, L. Psi tests with internally different machines. J Parapsychology, 1972, 36, 222-232.

Stanford, R.G. Towards reinterpreting psi events. J.A.S.P.R., 1978, 72, 197-214.

Stanford, R.G. & Palmer, J. Some statistical considerations concerning process-oriented research in parapsychology. J.A.S.P.R., 1972, 66, 166-179.

Tart, C.T., Palmer, J. & Redington, D.J. Effects of immediate feedback on ESP performance: a second study. J.A.S.P.R., 1979, 73, 151-165.

Timm, U. The measurement of psi. J.A.S.P.R., 1973, 67, 282-294.

Tippett, L.H.C. On the extreme individuals and the range of samples taken from a normal distribution. Biometrica, 1925, 17, 364-387.

Walker, E.H. Foundations of paraphysical and parapsychological phenomena. In Quantum physics and parapsychology. Laura Oteri (Ed.), Parapsychology Foundation, 1975.

Walker, E.H. Quantum mechanics,psi phenomena, the theory and suggestions for new experiments. J. of res. in psi phenomena, 1976, 1, 38-52.

West, D.J. Home testing ESP experiments: an examination of displacement effects. J.S.P.R., 1953, 37, 14-25.

West, D.J. Experimental parapsychology in Britain: a survey of recent work. J. Parapsychology, 1954, 18, 10-31.

West, D.J. Obituary G.W. Fisk. J.S.P.R., 1973, 47, 21-30.

West, D.J. & Fisk, G.W. A dual experiment with clock cards. J.S.P.R., 1953, 37, 185-197.

White, R.A. The limits of experimenter influence on psi test results: can any be set? J.A.S.P.R., 1976, 70, 333-369.

Wood, R., Kirk, J. & Braud, W. Free response GESP performance following ganzfeld stimulation vs. induced relaxation. E.J.P., 1977, 1, no. 4, 80-93.

LETTERS TO THE EDITOR

From D.J. Bierman: A methodological hint for research in RSPK cases.

One of the major tasks of the Poltergeist researcher is the registration of objective material on the case. To this purpose several times it has been tried to film or tape the phenomena on video, especially in the long lasting and high frequency of events type of cases. However with remarkably limited succes. Reports do suggest strongly that the phenomena stop whenever an objective registration device is operating. Remark that the phenomena are reported to be not very sensitive to extra witnesses or camera man as such. It is the device in operation which seems to be the essential element.
This elusive behaviour of the Poltergeist might be explained in terms of Observational Theories especially in terms of the Potential Observer Effect.
It is with this possible theoretical relevance in mind that I would like to suggest a slight modification of registration procedures in RSPK cases.
My suggestion is to use a series of short films from which a number has been exposed to light before filming by a co-experimenter who is also to be in charge of handing these films to the camera man in a random order. Random should be interpreted here as 'Unknownable for the camera man'.
One could try to compare real random with some preset algoritm (pseudo random).
The prediction is that phenomena will continue while filming with an already exposed film while they will stop while filming with an unused film although none of the people present on the site where the phenomena occur is aware of the condition of the film actually in use.
Any other outcome however has also some value with regard to evaluation of recent models in Parapsychology.

From Charles T. Tart:

In his review of the literature, "Physiological detectors of psi," Brian Millar, in reviewing a study of mine (Tart, 1963; 1975) states that

"There is one caveat, however, it seems that the analyses originally planned (Tart, 1975) were between experimental and control periods, rather than between each of these and control periods arbitrarily chosen 10 seconds after the start of each two-second trial. It is therefore not clear whether the three significances found in the 13 analyses were the results of psi or of data scrounging." (p. 459). He later mentions these same results as suspect in terms of "post-hocery." (p. 471).

Millar is correct in reporting that the main analysis planned was between experimental (shock) and "control" (non-shock) periods, but his speculation that the comparison of shock and non-shock periods with other control periods taken 10 seconds later was post-hoc and possibly biased is incorrect. Standard research procedure in psychophysiological research at the time of the study was to compare all experimental manipulations (shock and non-shock in this case) against as neutral a baseline period as possible. The defining of this neutral baseline as a period commencing 10 seconds after shock and non-shock trials was not done with any reference to empirical patterns in the data (as "data scrounging" and "post-hocery" imply), but was the result of an a priori decision that 10 seconds was probably sufficient time to allow for any physiological responses to the shock or non-shock trials to have largely or totally died out.

I find myself continually surprised that writers of review articles so frequently introduce speculations about studies under review without bothering to check with the authors of those studies before going into print, with consequences that detract from otherwise excellent articles."

From A.R.G. Owen, Editor New Horizons:

In his paper The Lloyd Effect published in your Journal 2(4), (1979) pp. 381-395, Brian Millar reports an experiment he performed specially to ascertain whether it is possible to obtain average evoked responses (AER) in the EEGs of "receiver" subjects to whom "sender" subjects are attempting to transmit signals telepathically. In its technical features Millar's experiment duplicates those of D.H. Lloyd (pseudonym) reported in New Horizons, 2, 1972, pp. 65-75, who claimed that although "because of lack of controls" the "experiment was not conclusive",

it did "directly suggest the capacity of the human brain to respond to psychic stimuli." Dr. "Lloyd" also concluded his paper, "Although the results as presented from this experiment are not conclusive, if further research does establish the validity of the application, then the AER to psychic stimuli will be the first truly objective primary correlate of psychic phenomenon, and will suggest the human brain does indeed respond to psychic phenomenon." I quote from Lloyd to indicate that his findings were put forward in a tentative and modest spirit appropriate to suggestions derived from what he described as a "pilot experiment."

Brian Millar's experiment appears to us as an excellent one and my associates and I are glad that it was done. We are however reluctant to follow him all the way in regard to the conclusions he draws from it. Millar obtained a negative result and therefore discusses two alternatives which he formulates as follows, "There are two possibilities, either his (Lloyd's) result depended upon some special percipient/agent/experimenter effect or it was artifactual." He goes on to express the opinion that the Lloyd effect was a statistical artifact. In our view he does not do full justice to the alternative. It seems from his own description that Millar's subjects were not selected on the basis of having previously given evidence of psi ability (whether statistically significant or not). Instead they were students, randomly acquired in respect of ESP, who themselves recruited the required complement of additional subjects. Thus the comparison between his and the Lloyd experiment is not fully cogent because it fails to compare like with like. Lloyd's subjects consisted of people who, in spontaneous situations or informal ESP experiments, had all shown at least some prima facie evidence of psi activity. This remark of course does not detract from the importance and value of Millar's experiment. For instance his result may exemplify what, indeed, might be true--that telepathy does not occur even unconsciously among randomly chosen subjects.

However even this last inference cannot yet be inferred with certainty. Besides an ability for telepathic transmission a further capacity is required of a successful sender. As was said in New Horizons, 1 (5), 1975, page 229, "Dr. Lloyd had difficulty finding good senders because it requires a high degree of concentration to be able to send a stimulus on a repetitive basis . . . senders do require a considerable amount of training and it seems that some people never

become proficient at it." Earlier in the same article the rapporteur (Dr. J.C. Whitton) says "Sometimes . . . the synchronization fails--the sender says he's losing the ability to keep up with the equipment." Just what is the best way of deciding whether a sender is efficient (other than by his own impression as to whether he can "keep up") is, admittedly, not clear. Until this problem has been reliably solved however, some doubt will remain as to the interpretation of even the most impeccably conducted experiments.

Dr. Lloyd and his co-workers readily concede the point that they did not attempt to perform a test of statistical significance on the effect they perceived. They made no pretence of so doing. However it is not the case that they "dredged" their data, selecting only tracings which suited their thesis. It might even be that a careful re-examination using sophisticated statistical tools of the kind not customarily used in the audiometric field in 1969 when the Lloyd research began, would indicate statistical significance. They grant also that the results obtained in 1974 in the only repetition of the Lloyd experiment which they have attempted though encouraging were not conclusive because the subjects' waveforms differed from those of the (different) subjects in the 1969 experiment. However it is only continued experimentation that can actually resolve some of the questions still at issue. Brian Millar's experiment is a most valuable and important one, as is his analysis of his results. We are very pleased that it was undertaken and published. We would only suggest that there are some unresolved problems remaining and that still further work of high sophistication is needed before the hypothesis of evoked response can be either validated or laid to rest.

From B. Millar: Rejoinder to Dr. Tart

I am most grateful to Dr. Tart for his correction of my review. It appears that the indiscretion in his study was not post-hocery but was instead selective reporting. The 7 analyses between shock and non-shock which Tart now reports to have been the main ones were not so much as mentioned in the original paper. The 13 analyses reported were, it seems, secondary but planned.

The practice of omitting to report planned analyses which

did not turn out is a most misleading one. If enough analyses are performed some "significances" will invariably be found even in random numbers. This procedure is unfortunately endemic in psychology (Barber, 1976, p. 21-22), where it has been estimated by Neher (1967) that one analysis from three performed is reported by the average experimenter. Parapsychology particularly cannot afford to waste its slim resources in exploring the blind alleys to which such a practice gives rise.

The current example only came to light 12 years after the publication of the paper. Were there yet other analyses which still remain unreported? Assuming there are not, 3 analyses from a total of 20 are significant at or better than the 5% level. (While I have been unable to understand the application of the Friedman method to finger pulse volume data, I assume in the following that it is correct.) How likely is it that such a state of affairs would occur by chance alone? An exact answer can be calculated only if the intercorrelations are known; however an approximate estimate can easily be made on the assumption that the analyses are substantially independent. The binomial probability of finding at least 3 analyses "significant" from 20 is 0.075, a result which is not itself significant.

In spite of this outcome it is reasonable to ask if anything in the data speaks against a chance explanation. It is noteworthy that the EEG significances are both shock and non-shock complexity measures. Furthermore, save for one exception (GSR, shock) all dependent measures for both types of experimental period are consistent with a pattern of increased arousal as compared to control periods. However, it remains true that when allowance is made for the unreported analyses the overall evidence for the presence of psi in the experiment is not inconsiderably attenuated. Perhaps some real psi effects drowned in a sea of multiple analyses: only further experiment can decide. Statistical sharpening techniques such as the "Edinburgh split" in which predictor and confirmation data are acquired in the same experiment are useful in avoiding this kind of situation, where detailed predictions cannot be made on a priori grounds.

I have lamented (p.469) the singularly poor reporting of work in this area. In addition to the major problem of selective reporting discussed above for Tart's study, it should perhaps be mentioned that important matters such as the source of the random selection of shock/non-shock targets, as well as whether the polygraph measures were

scored blind on the basis of pre-determined criteria, are not adequately covered in the paper.

I agree with Dr. Tart that a reviewer should ideally query every hiatus in a report, with the author. In practice, however, this goal is somewhat utopian. The amount of work involved would, alas, rapidly assume Herculean proportions. The best one can do is to follow up a few papers in fine detail. Consequently I fear that the onus for adequate dissemination of results must fall where it has always fallen, namely square on the writers of research reports.

REFERENCES

Barber, T.X., Pitfalls in human research: ten pivotal points. New York, Pergamon, 1976.

Neher, A. Probability pyramiding, research error and the need for independent replication. Psychological Record, 1967, 17, 257-262.

ERRATUM

The following references were inadvertently omitted from the article "Physiological detectors of psi" by Brian Millar, in the preceding EJP, 1979, 2, no. 4, 456-478.

American Society for Psychical Research, Research Committe, Report of the research committee for 1958. J.A.S.P.R., 1959, 53, 69-71.

Duane, T.D., Behrendt, T., Extrasensory electroencephalograp hic induction between identical twins. Science, 1965, 150, 367.

Esser, A.H., Etter, T.L., Chamberlain, W.B., Preliminary report: physiological concomitants of "communication" between isolated subjects. Int.J. Parapsychology, 1967, 9, 53-56.

Figar, S. The application of plethysmography to the objective study of so-called extrasensory perception. J.S.P.R., 1959, 40, 162-172.

Kennedy, J.E., Taddonio, J.L. Experimenter effects in parapsychological research. J. Parapsychology, 1976, 40, 1-33.

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 3

NUMBER 2

MAY 1980

Robert L. Morris
Psychology Department
University of Edinburgh

EDITORIAL STAFF

Martin Johnson
Sybo A. Schouten

REGIONAL EDITORS

U.K.
Dr. John Beloff
University of Edinburgh
Department of Psychology
7 George Square
Edinburgh EH8 9JZ

Sweden
Dr. Rolf Ejvegaard
Mellansjovagen 35
S-14146 Huddinge

U.S.A.
Ms. Rhea White
2 Plane Tree Lane
Dix Hills
New York 11746

Australia
Dr. H.H.J. Keil
University of Tasmania
Box 252 C, G.P.O.
Hobart, Tasmania

W. Germany
Prof. Dr. Hans Bender
Prof. Dr. Johann Mischo
Institut fur Grenzgebiete der
Psychologie und Psychohygiene
78 Freiburg i. Br.
Eichhalde 12

Iceland
Dr. Erlendur Haraldsson
University of Iceland
Department of Psychology
Reykjavik

Italy
Dr. P. Cassoli
39, Via Valeriani
I-40134 Bologna

Prof. Dr. Emilio Servadio
Via di Villa Emiliani, 4-int.7
I-00197 Roma

The European Journal of Parapsychology publishes original experimental and theoretical papers on parapsychology. In addition, articles presenting reviews of literature, case studies and papers on related topics are accepted in so far as the subject can be considered relevant to the understanding of parapsychological phenomena or to the methodology of empirical research in parapsychology. Translations of papers originally published in a language other than English are also welcome.

Manuscripts should be written in English and should be typed, double-spaced throughout, with a wide margin, on one side of the page only. The organization of the paper must be clearly indicated by appropriate headings and subheadings. Include names of authors, positions and other affiliations, and mailing addresses. Each table

(continued inside back cover)

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

CONTENTS

Johnson, M.U.	J.B. Rhine	p.119
Blackmore, S.J.	Correlation between ESP and memory	p.127
Braud, W.G.	Project Chicken Little: A precognition experiment involving the SKYLAB space station	p.149
Gruber, E.R.	PK effects on pre-recorded group behavior of living systems	p.167
Keil, H.H.J.	The distribution of psi ability and psi performance	p.177
Locke, R.G.	Communication and imagery resonance: A theoretical approach to psychic healing	p.185
Alvarado, C.S.	On the 'transference' of psychic abilities: A historical note	p.209
	Call for papers	p.212

PUBLICATION POLICY

Twice a year the Parapsychology Laboratory of the University of Utrecht publishes the European Journal of Parapsychology. The object of the European Journal of Parapsychology is to stimulate and enhance the activity in this field, especially in our corner of the world, by communicating research results and issues related to professional parapsychology. Although there will be an emphasis on experimental work, theoretical articles are also welcome. Contributions from all over the world will appear in the journal.

A hallmark of the European Journal of Parapsychology is the attempt to avoid selective reporting, that is, the tendency to bury 'negative' results and only to publish studies that 'turn out'. To avoid turning the journal into a graveyard for all 'unsuccessful' studies, we require that the acceptance or rejection of a manuscript should take place prior to the phase when the experimental data are collected. The quality of the design and methodology and the rationale of the study are considered more important than the level of significance of the outcome of the study. As a practical rule, we advise the potential contributor of an article to submit the design of his planned study before the work is actually carried out. The rationale of the study should be stated, as well as all the hypotheses related to it. Furthermore one should try to specify the number of subjects, the number of trials, etc., plus the type of statistical methods one plans to use for evaluation.

Priority will be given to the publication of studies which fulfil the above-stated publication policy.

The final manuscript with presentation of results must reach us two months in advance of the official publication dates, which are May 1st and November 1st.

J.B. RHINE

1895-1980

J.B.Rhine, the nestor of parapsychology and by any measure the most influencial of all parapsychologists so far, died suddenly on the morning of February 20th, in his home in Hillsboro, Orange County, North Carolina. His wife and research-companion Dr Louisa Rhine, three daughters and one son survive him.

J.B.Rhine's contribution to parapsychology was complex, diversified and wide-ranging. Dr J.B.Rhine's illustrious career illustrates and substantiates to a great degree the concept recognized as the American Dream - the dream of being a frontier-man, working hard, showing personal initiative and being successful.

J.B.Rhine was born on September 29th 1895, in Juniata County, Pennsylvania. He learnt early to be industrious and to pave his own way. He was educated at Ohio Northern University, the College of Wooster and the University of Chicago where he received his M.A. (1923) and a Ph.D. from the Department of Botany in 1925. The main features of his further long and exceptionally active life should be rather well-known to most readers of a parapsychological journal. Perhaps the most important milestone in his life was reached in 1920, when he married Louisa E. Weckesser, his constant support in life and fellow-researcher. In 1926, the Rhines, somewhat influenced by a talk on spiritism, delivered by no less a person than Sir Conan Doyle, went on their epic tour to Harvard to see Professor William McDougall who was just about to go to Europe for a longer period of time. A year later Professor McDougall accepted the offered Chairmanship of the then created Department of Psychology at Duke University. In September 1927 the Rhines came to Duke University upon the invitation of Professor McDougall. They gladly accepted the chance for post-doctoral

studies in psychology and the area of psychical research since Professor McDougall was the person best prepared at that time to serve as their mentor both in psychology and 'psychical research'.

The first challenge was to investigate mediumship and claims of spiritual survival. J.B.Rhine soon came to the opinion that the claims, at that stage of research ,could neither be verified nor entirely falsified. Only extended knowledge about paranormal communication between the living and discoveries about certain boundary conditions could lead to a more meaningful approach towards the survival issue.

With tremendous energy, J.B.Rhine then broadened his field of study to include a wide range of paranormal phenomena. He made a great number of methodological improvements; among others he made extensive use of statistical methods to carefully test the hypotheses of telepathy and clairvoyance. He pointed out that tests considered as tests of telepathy were not really tests of telepathy. Positive results could just as well be interpreted in terms of clairvoyance. One of his great conceptual contributions was the design and introduction of what was considered as a test of pure clairvoyance, by means of the well-known Down-Through cardcalling technique.

As early as December 1933 the first known laboratory test of prophecy (precognition) was also carried out at Duke. Rhine had again outlined a surprisingly simple procedure by which the hypothesis of the mind's independence of time could be tested according to scientific criteria. Almost simultaneously (in early 1934) Rhine carried out a test of 'direct action of mind over matter', described by Rhine as a test of psychokinesis. Dr. Rhine was not the first to use statistical methods in parapsychology, but the first to make extensive use of the approach in a coherent, planned research program. Rhine was not the inventor of the term parapsychology but the first one to professionalize its use. The same can be said about the term extrasensory perception.

In 1934 the time was ripe for the publication of 'EXTRASENSORY PERCEPTION', a book which has already a recognized place in the history of ideas, independent of what may happen to parapsychology as a science. This book is, according to Michael McVaugh and Seymour H.Mauskopf, historians of science, something of a paradigmatic work for parapsychology, in that it synthesizes previous work and sets a pattern for subsequent experimental research. Also, thanks to the research work published in this book and based on how the environment

reacted to it, a special research center for parapsychology was set up at Duke. It was given the name: the Parapsychology Laboratory of Duke University. J.B.Rhine was appointed its director. He stayed in this position for 28 years, i.e. until his mandatory retirement age as professor of psychology.

Three years after the publication of 'EXTRASENSORY PERCEPTION', and the extended, heated debate that followed, J.B.Rhine founded the Journal of Parapsychology with the help of Professor McDougall. This journal has been strongly influenced by Dr. Rhine's ideas of what constitutes the right publication policy. It has very strongly reflected his esteem of parapsychology and his ideal of science. Needless to say, the journal has developed as perhaps the leading journal in experimental parapsychology.

Rhine has a very extensive production. If one counts all his professional papers, published speeches, articles and editorials it amounts to several hundred publications. Another 'classic' is 'EXTRASENSORY PERCEPTION AFTER SIXTY YEARS' (1940) published jointly with four colleagues, with J.B.Rhine as spiritus rector. Among more professional books, one should also mention the textbook: 'PARAPSYCHOLOGY: FRONTIER SCIENCE OF THE MIND', which he wrote together with J. Gaither Pratt. Rhine also wrote a number of books for the general public: 'NEW FRONTIERS OF THE MIND' (1937) - a book-of-the-month (1937); 'THE REACH OF THE MIND' (1947); 'NEW WORLD OF THE MIND' (1953); 'PARAPSYCHOLOGY: FROM DUKE TO FRNM', (1965) (Jointly with associates). Together with Robert Brier he edited 'PARAPSYCHOLOGY TODAY' (1968) and in 1971 he edited 'PROGRESS IN PARAPSYCHOLOGY'. Dr.Rhine was also a very much sought after lecturer. I am not sure I could properly characterize him as a very good speaker, his voice was not very suitable for large audiences, but in some way this limitation made one very attentive.

Rhine always maintained a sober style in relation to the mass-media. He did not fall for the temptation to sensationalize research, either at the Parapsychology Laboratory, or later on at the FRNM. He was very well aware that over-selling could backfire. His way of handling, at a rather advanced age, the very sensitive and unpleasant Levy affair was admirable. He took responsibility, and became respected.

The educational endeavours of J.B.Rhine were very extensive. The majority of those who work full-time in parapsychology today have had shorter or longer stays with Rhine. Many of them have obtained their basic training either at Duke or at the FRNM, under Rhine's

leadership.

J.B.Rhine had an extremely wide-ranging correspondence. He stood in correspondence with almost all who took a more or less professional interest in the field. He was a master as a letter-writer. The main bulk of his correspondences have to the best of my knowledge been donated to Duke University. Many Americans as well as non-Americans were brought into our field or into closer contact with the field through J.B.Rhine. Of considerable educational value were the 'review-meetings' that the FRNM arranged in the late 60's and early 70's. The system of inviting (usually a hand-picked) visiting research fellow for a period of work and training has certainly had an impact on parapsychology. The idea has been adopted by our laboratory, where one of the positions is intended for a visiting research associate.

J.B.Rhine should also be commemorated as the founder of the Parapsychological Association, the only internationally recognized organisation of professional parapsychologists in the world. The PA today plays an instrumental role in maintaining professional and ethical standards in our field and in the reporting of research findings in parapsychology.

What was J.B.Rhine's main contribution to parapsychology? As Brian Mackenzie, expert on the history of psychology has stated, Rhine brought about a change from the amateuristic approach executed by the private societies towards professionalism. "With Rhine as a focal point it developed its own values, standards, procedures, problems, language and audience. Problem selection was initially based upon the choice of the best definable issues, and thereafter largely upon the results of immediately previous research. The isolation of 'extra-sensory perception' (a tern which Rhine coined) as an area for concentration amongst all paranormal problems and phenomena which were topics of popular and amateur interest, was the first important result of such professional restriction of attention" (J.P., Vol.42, 3, p.206). Rhine's long life was a fierce struggle for an idea. In Rhine's opinion the objective of parapsychology should be to uncover the hidden dimensions of man. Through this research process, the inner, true nature of man would be discovered and this discovery will help him to liberate himself from the bonds of materialism.

Hence, we can see that Rhine, like the founders of the societies, had extra-scientific motives for the scientific endeavours. Regarding the research method, Rhine seems to have been a realist and a naive verificationist. His ideal of science was in many respects congruous

with that of the natural sciences at about the turn of the century. As has often been stated, metaphysics is a necessary precedent and ingredient in all theoretical systems, and especially in the early phases of a science "the importance of the world-picture is greater, as many preconceptions of the territory still have to be articulated" (Nilsson,I., EJP, Vol. 1, November 1975, p.53).

In short, J.B.Rhine thought that the findings of parapsychology strongly challenged the world-view of the materialist. He even played with the idea that the realities which might be revealed by parapsychologists might in a way validate some of the basic concepts in most religious systems and thus give rise to a 'natural theology' on a scientific and empirical footing. He firmly believed that parapsychology had a bearing on ethical and political issues, and, "Like the humanistic psychologists in the early 1950's, Rhine found Western man in an existential crisis, religion fading away, but materialism and communism gaining in strength" (Nilsson, I., Ibid). As already emphasized, he thought that parapsychology could create a 'natural theology', based on empirical facts about 'man's nature' instead of faith based on dogma. This, his general position, seems to me to be highlighted by the words his wife, Louisa E.Rhine uses in her dedication of her book: 'ESP IN LIFE AND LAB':

> To JBR
> Whose urge to understand the uniqueness of
> man's nature would not be satisfied by dogma
> or speculation, but only by the results of
> objective scientific investigation.

Rhine as a philosopher remains, admittedly, a quite controversial matter. After all, he did not have any training as a philosopher. About his burning interest in understanding Man and his place in the Universe however, no doubts could be had.

It is almost impossible to think of J.B.Rhine without also thinking of his extra-ordinary wife, life-companion and fellow-researcher, Mrs. Louisa E.Rhine. Especially in his advanced years he was always keen to mention the crucial role she had played in his career and thinking. J.B.Rhine has already been the subject of several papers and even thesis work. If parapsychology is distinctively going to make its case as a science, I predict that future investigations will also focus on the role Louisa E.Rhine has played in his thinking and work.

I have not infrequently come across the allegation that Rhine

neglected field-studies and the spontaneous cases. This allegation can to a substantial degree be refuted by the fact that he encouraged his wife to devote much of her work to the collection of spontaneous cases. As a matter of fact she turned into one of the world's experts in this area. He never denied the heuristic values of field-studies and the study of spontaneous cases, but remained hesistant about how to evaluate such observations in a strictly scientific way.

I think it is correct to state that J.B.Rhine felt a great deal of frustration about the slow progress in parapsychology. He had hoped to see the birth of a truly 'repeatable experiment' in parapsychology. In the early seventies the anpsi experiments at his Institute and elsewhere seemed to promise that a breakthrough was imminent. His great enthusiasm for the anpsi approach may be explicable partly because the results looked extremely promising, the approach had a rather good rationale, and finally since J.B.Rhine obtained his basic training as a scientist in the field of biology. It is no secret that he worked hard and published hard - perhaps too hard for such a breaktrough. After the incident in 1974, he took a more resigned position, and from our correspondences and talks I got the impression that he felt that parapsychology found itself in a somewhat messy situation - a situation which probably could not be cured by the new conceptual tools which 'observational theories' offered, even if wrapped in a more or less elegant mathematical formalism. One could assume that his lack of interest in the 'observational theories' entirely depended on the fact that he could not possibly grasp all their technicalities. It may have been a contributary factor, but it certainly is not a sufficient explanation.

Like most other exceptional persons, J.B.Rhine remained a somewhat controversial person. Ideas about him ranged from seeing in him a new Columbus or Copernicus to viewing him as a somewhat restricted scientist, being obsessed by the execution of tedious and repetitive card-calling and dice-throwing experiments. He knew that he was a controversial person, and he did not complain about it. He pursued his life in 'his way'. He was also fully aware that even during his lifetime he had become a legend. He knew that he belonged to the history of science and ideas. J.B.Rhine's personality was complex and diversified. I certainly do not pretend to be able to give a complete description of Dr. Rhine's personality, nor do I wish to try to do that. However, I would like to give a few personal impressions. He was a strikingly charismatic person, he possessed great ability to influence other people. The essence of his charm may have been the unique blend of a somewhat shy, smiling country boy and the

iron-willed leader with a military background, backed up by unusual intellectual ability and a rich life-experience. I recall that a person who worked with him for a considerable time characterized him as "not as much a person as a force of nature". That statement reflects of course a metaphore and an idiosyncracy. However, it does say something about J.B.Rhine as a researchleader and administrator.

To say the least, he was not always easy to work with. If he once had taken a position, it was not easy to change his view. He expected much from his staff-members as regards dedication and hard work, but he expected still more from himself. It is telling, that even Christmas Eves were dedicated to work on behalf of parapsychology. J.B.Rhine could be tough and unsentimental. The other side of the coin is that he could be very warm and compassionate in interpersonal affairs. I just recall an episode when he had to bring bad news, the news about the death of one of the parents of one of the visiting research associates at his laboratory. He tried, but became successively so deeply moved that he could not fulfill the mission. He had to ask his wife for help.

As a scientist Dr. Rhine was respected, even among his opponents, for reliability and honesty. As far as I can recall, only two people have publically questioned this, and both have had to apologize for their unfounded allegations. To me, J.B.Rhine represented a great father figure. Without J.B.Rhine, I doubt that I would have become a parapsychologist.

One should finally mention that Rhine was a great admirer of 'Mother Nature'. He considered it almost a perversity to live in a city. Especially on the occasions when we walked in his wood, relaxed in the beautiful Duke University Garden, or sipped his home-made cider on his farm, I felt a strong feeling of 'rapport'.

To be able to see J.B.Rhine in his right perspective, we must surely beg some time. Firstly we must await the further development of parapsychology as a science. Not the least the turmoil of sentiment must be given a chance to settle.

We are not simply mourning "another distinguished member of the PA", but essentially its founder. I hope time will permit that the PA shall be able to pay proper tribute to the memory of J.B.Rhine at the forthcoming PA Convention in Reykjavik and at those that will surely follow.

Meanwhile, let all of us extend our deep compassion and sympathy towards those closest to him, that remain.

Martin Johnson

CORRELATIONS BETWEEN ESP AND MEMORY

S.J. Blackmore

University of Utrecht

I have discussed elsewhere the importance of attempts to integrate ESP into psychology (Blackmore, 1980). One of the first questions to ask may be whether ESP resembles other cognitive processes such as memory and perception. There are several ways in which this may be investigated. They include 1. The study of errors made in ESP, 2. Correlations between ESP and memory or other cognitive skills, 3. Studies of the effects of varying target material, 4. The use of the recall situation as a psi-conducive state and 5. The study of the effects of associative habits on ESP. Some of these approaches have already been discussed. This paper will consider only one of these, that is the correlations between ESP and memory.

Modern research on ESP and memory began after the publication of Roll's 'memory theory of ESP' (Roll, 1966). This was based on the idea that "the ESP response consists of the percipient's own memory traces and that the effect of the external (ESP) stimulus is to activate memory traces rather than supply new ideas or images". Roll drew several predictions from this theory among them that "If memory traces are vehicles for psi impressions, we expect a person with good recall ability to perform well in ESP and, conversely, a good ESP subject to perform well in memory tasks." (Roll 1966, p.510). This prediction is relatively easy to test by correlating performance on ESP and memory tasks and this has been done. But before considering the research that has been carried out I wish to consider what conclusions might be drawn if such a correlation were found.

Firstly Roll predicted a positive correlation between recall and ESP on the basis of his 'memory theory of ESP', but I do not believe this necessarily follows from that theory. If the role of the ESP stimulus is to revive, and in so doing to select, the appropriate memory trace then the major task of recall is carried out paranormally and there is no particular reason to suppose that the same person will be good at normal recall and this other paranormal process. On a two-process theory of memory we may argue that Roll's theory involves paranormal retrieval but normal recognition and one might then predict a positive correlation between ESP and measures of recognition, but not recall. While this is not the only possible interpretation of Roll's theory it does indicate that a positive or negative correlation between ESP and recall scores is not necessarily evidence either for or against it. There are several other models which relate ESP and memory. For example Carington's 'association theory of telepathy' suggests that associations between ideas and images (psychons) tend to persist independently of the person who initiated them and may be utilised by others so leading to the occurrence of telepathy. Although this theory has many drawbacks and can only account for telepathy, not clairvoyance, it treats memory and telepathy as equivalent processes and so would, I believe, predict a positive correlation between them. The same may be said of Price's 'Psychic ether hypothesis' (Price, 1939), Marshall's physical theory of ESP and memory (Marshall, 1960) and what Roll (1966) calls the psi-trace theory of memory, that is that the brain acts as a kind of token object in memory, allowing for retrieval of memories stored as psi-traces. Since in all these theories ESP and memory are seen as the same process I believe they might all predict a positive correlation between ESP and memory.

Rao, Morrison and Davis (1977) argued for a negative correlation on the basis of Bergson's idea that the brain's selective processes act as a protection against recalling and experiencing too much. If ESP is seen as the result of errors escaping cortical surveillance then better ESP would occur with a less efficient memory. But even this is not a simple deduction. Bergson (1911) attributed a selective process to the nervous system, suggesting that in remembering it selects the desired information from all that available. If all information is potentially available then random escapes through the surveillance would not lead to ESP but to confusion. For correct ESP some sort of selection is still needed. Therefore a better selective process may be able to select ESP information better as well as memory. This would lead to the opposite prediction to that made by Rao and his associates.

If a consistent correlation between ESP and memory were found it would not be obvious which theories of ESP would be supported or refuted. Indeed I do not think there is any theory of ESP so clearly formulated that it would allow definite conclusions to be drawn. Nonetheless, such a finding could be useful. It would indicate whether there was a similarity between the processes involved in ESP and memory, or in the factors which affect them, and this could provide a start.

We may now consider the existing evidence for such a correlation. Feather (1965; 1967) was the first to report a correlation between ESP and memory. She gave subjects two runs of clairvoyance, a memory test for a list of 25 ESP symbols and then two more runs of clairvoyance, and found a positive correlation between recall and ESP scores. This study was complicated by the fact that the memory test was very hard, with some subjects scoring below chance on it. Rao et al (1977) suggested that ESP may have operated in the memory test as well so complicating any conclusions drawn, since the correlation could have been due to ESP operating in both tests rather than to any relationship between ESP and memory.

Many subsequent studies have tested the correlation between ESP and memory but mostly only as a secondary or even post hoc analysis. For example Stanford (1970) tested several hypotheses in one experiment. Subjects listened to a story followed by a set of multiple choice questions. The 'correct' answer to these questions was randomly determined. For some the answer was given in the story, for some it was implied and for others it was not mentioned. According to the negative response bias hypothesis one would expect that higher ESP scores would be obtained for those questions which were included in the story (when the 'correct' answer was counter-story) than for those implied or not mentioned, because memory for the story would produce a bias against them. In addition it was expected that subjects with a better memory would show a greater effect. 30 subjects completed the ESP test, a test of incidental learning for the details of a room and various other tasks, not relevant here.

The results confirmed the negative response bias hypothesis. When the answer was specified in the story other answers were more often correct (that is by ESP) than when they were not specified. Also, when the answer had been given, there were few counter-story responses when the 'correct' answer was the same as that given, but more when it was different. High memory subjects obtained significant scores on all counter-story responses, while low memory subjects scored in the

predicted direction but not significantly. Of most importance here though is that on the crucial counter-story responses the high memory subjects scored higher than the low memory subjects, but when no answer was given or implied in the story there was no difference. This led Stanford to conclude that memory, per se, is not related to ESP performance, but the effect is indirect in terms of its relationship to response bias. Thus although Stanford found a positive relationship between ESP and memory he attributed it to an indirect effect of memory.

In a simpler experiment Peterson (1972) gave 28 subjects both memory and clairvoyance tasks using two different kinds of ESP symbol but found no significant relationship between ESP and memory scores. Kanthamani and Rao (1975), developed a method for studying ESP and memory which has been more widely adopted than any other. They argued that previous studies had measured ESP and memory consecutively. They wanted to study their interaction trial by trial, bringing the two as close together as possible. The original method consisted of giving subjects a memory task and incorporating an ESP test into the recall phase. Their hypothesis that ESP would occur differently on recall-correct and recall-wrong trials was tested, but will not be discussed here. I have previously criticised the rationale behind this method (Blackmore, 1980). More relevant here is that an ESP and a recall score was obtained for each subject. Kanthamani and Rao pooled data from 62 subjects and 4 series, correlated these scores and obtained a significant positive correlation of 0.284 (They give an associated t value of 2.876, $p<.01$, 2-tailed. My calculations for their value of r give $t=2.294$, $p=0.025$, 2-tailed. This still shows a significant positive correlation). However, this finding should be treated with caution since the analysis was conducted post hoc, and the data from four dissimilar experiments were pooled.

Parker (1976) tested 40 subjects' digit span in an immediate memory task consisting of progressive learning of lists of 3 to 9 digits. An incorporated ESP task consisted of placing the digit answer on one of two lines. Although overall ESP scores were at chance a significant negative correlation between ESP and memory scores was obtained ($r=-0.347$, $t=2.28$, $p<.05$). In a second similar series a non-significant positive correlation ($r=0.107$) was obtained.

Rao, Morrison and Davis (1977) carried out two sets of experiments in which subjects memorised lists of paired associates, of a trigram and a word. Some ostensibly recall trials were actually ESP trials and the results of each type were correlated. In the first set of seven

experiments ESP scores were close to chance as were the correlations with memory scores and there was no apparent consistency in the direction of the correlations. In the second set, comprising seven further experiments, data for 118 subjects were pooled and (post hoc) a negative correlation between ESP and memory scores obtained ($r=-0.18$, $t=1.99$, $p<.05$, 2-tailed).

Finally in a similar experiment Rao (1978) tested both ESP and memory for 25 pairs of a trigram and a word. 93 subjects were tested in three groups and a small positive correlation between memory and ESP scores obtained ($r=0.18$, $t=1.70$). In addition the data fell into two groups, the first being only computer scored and the second hand scored first. It was the data from the second group which contributed most to the correlation. Rao discussed this difference in terms of a possible checker effect or differences in the subject-experimenter relationship. He also pointed out "Thus, the research has come full circle: no significant correlation in the exploratory study, significant negative correlation in the replication study, and now a suggestive positive correlation for overall results and a strong positive correlation for group 2." (Rao 1978, p.176).

Whether or not one considers the correlation 'suggestive', these results clearly show that there is no simple answer to the question of whether memory and ESP ability are correlated. Indeed the results of Rao and his colleagues confirm the confusing pattern seen in the previous results, that is, a positive correlation found by Feather (1965; 1967), no correlation found by Peterson (1972) and a negative correlation found by Parker (1976).

What is one to make of these results? One could conclude simply that there is no correlation between the two abilities and that the small correlations obtained were spurious. Alternatively the correlations could be valid, but the results obscured by other variables. Perhaps most important of these variables to consider is the type of memory task used. Feather's very difficult task of learning 25 ESP symbols in 15-20 seconds bears little resemblance to learning paired associates or testing digit span. These task differences could be the cause of the varied results, although there must then be some other explanation for the fact that Rao and his colleagues obtained very varied results using the same memory task. But ignoring the latter problem for the moment, what task variables might be important?

One possibility is task difficulty. If ESP can be compared with remembering it is like very difficult remembering. Therefore a harder

memory task may provide a better correlation. In this context it is interesting to note that Feather's task was very difficult and she obtained a positive correlation. This suggestion cannot account for Rao's variable results but nonetheless it could easily be tested.

A second possibility is that a similar ESP and memory task is required. Although this is hard to measure it seems that those studies which used apparently most nearly comparable tasks did not produce the most consistent correlations. These are Feather (1965; 1967), Peterson (1972) and Rao, Morrison and Davis (1977).

One final difference between the tasks may be whether they test primary or secondary memory. Irwin (1979) argued that Roll's theory only predicts a positive correlation between ESP and measures of secondary memory, and that there might be no correlation or a negative correlation with tests of primary memory. The difference lies in the ways in which information is stored. In primary memory it is less deeply processed and represented in such a way that it can only be retrieved in the short term. Information in secondary memory has typically undergone detailed semantic processing and may be available indefinitely. Irwin argued that both Feather and Kanthamani and Rao used secondary memory tasks and obtained positive correlations with ESP, while Parker tested primary memory and obtained a negative correlation. The problem then is the varied results obtained by Rao et al (1977) and Rao (1978). Irwin suggested that their task permitted use of both types of memory, in which case the varied results might depend on subject strategies. This suggests that better correlations would be obtained when subjects were forced to depend on one type of strategy only. This could undoubtedly be tested. Since the available data does not make it possible to conclude whether or not there is a correlation between ESP and memory, more research is clearly required.

The experiments reported here were intended to test further the correlation between ESP and memory scores and to provide more direct tests of the various theories of ESP and memory. It was argued that one interpretation of Roll's theory predicts a positive correlation between scores on ESP and tests of recognition, but not of recall. A first series of experiments therefore included tests of both recall and recognition for a word list and correlated performance with ESP scores. Since the recall-recognition ratio has been used as a measure of the superiority of recognition over retrieval this was also measured. In an attempt to explore other memory tasks another series of experiments included a test of incidental learning. It has been argued that if ESP is like remembering it is most like very difficult

remembering or remembering for which one has had no opportunity to learn. A test of incidental learning is perhaps most closely comparable to this situation and may provide a better correlation with ESP. Also if ESP is like difficult remembering then we may expect better correlations when the memory task is particularly hard. An experiment was designed to compare the correlations on two memory tasks differing in difficulty.

In view of the previous discussion on types of ESP task, we may argue that a positive correlation can only be expected where the ESP task and the memory task are, as far as that is possible, comparable. This has certainly not usually been the case, although the studies of Feather (1965) and Peterson (1972) are possible exceptions and Rao et al (1977) used a comparable ESP and paired-associate learning task. The fact that these studies did not provide strong evidence for a correlation may indicate that this approach is unlikely to be successful although it does remain a possibility. Finally Irwin (1979) suggested that the crucial difference may be whether the memory task measures primary or secondary memory. If this hypothesis had been published prior to these experiments being carried out it would have been tested directly. Nonetheless some analysis is possible to determine whether this hypothesis is compatible with the data obtained here and further experiments are planned.

In all these experiments the simplest prediction is that according to any of the memory theories a positive correlation between ESP and memory scores is expected. Other theories do not predict such a correlation but are not incompatible with it. To distinguish between them other approaches, perhaps along the lines mentioned above, will be needed. But most important is that the results might show whether there is any similarity between the processes of ESP and memory, or the factors which affect them.

The results of 6 experiments are reported here. The preliminary experiments were carried out with the intention of finding the best methods to use in later experiments. Since they suffer from methodological weaknesses they are only reported in outline here (for further details see Blackmore, 1980).

PRELIMINARY EXPERIMENTS

EXPERIMENT 1

20 students in a parapsychology class took part in two sessions each consisting of 100 trials of a standard clairvoyance task with ESP cards and a memory test in three parts. First Ss were shown a list of 25 common nouns, on an overhead projector, for four minutes, and given a further four minutes in which to recall them. This was followed by a second recall test with 25 new words and finally Ss were given a list of 125 words in which the first set of 25 was embedded and were asked to underline any which they thought had occurred in the previous test. The number of words each S correctly recalled and recognised was recorded and the recall recognition ratio for each calculated. These scores were then correlated with ESP scores.

Results

Each S was given a combined ESP score for the two sessions. The mean score was 41.1 which is not significantly different from MCE of 40.0 ($t=1.12$, 19df, $p=0.26$).

The memory tests were effective in that scores varied widely. No S recalled all 50 words correctly and only 2 Ss recognised the 25 correctly. The mean number of words recalled was 37.7 and recognised 21.65. Mean recall-recognition ratio was 0.16. When correlated with ESP scores the following results were obtained:

Recall	$r= 0.09$	$t= 0.38$	$p= 0.71$	19 df
Recognition	$r= 0.002$	$t= 0.01$	$p= 0.99$	19 df
Ratio	$r= 0.21$	$t= 0.90$	$p= 0.38$	19 df

None of these correlations is significant.

Discussion

There were two main faults in the design of this experiment. Firstly one target order was used for all Ss, so allowing for a stacking effect to occur. Individual target orders were used in later

experiments. Secondly, the ESP tests were conducted prior to the memory tests and so Ss already knew their ESP scores. Conceivably this could lead to a spurious correlation between the two. This possibility was eliminated in later experiments. Also the separation of the tests may reduce any genuine correlation if variables such as mood, time of day, or long term changes of some kind affect either ESP or memory scores. Ideally the tests should be carried out at the same time. This was done in experiment 6. A final problem is the small number of subjects. The experiment was repeated with more Ss.

Apart from these faults in design a problem arises in the attempt to correlate one measure, memory test scores, with ESP scores since the latter provide no independent evidence that any ESP was occurring. Two arguments are applicable here.

1. It is often argued in parapsychology that when overall ESP scores are close to chance the effects of ESP may be masked by the opposing tendencies of psi-hitting and psi-missing. If these can be separated then the effects of ESP can be seen. This has been achieved for example by distinguishing groups of Ss according to belief (Schmeidler, 1945) or by various personality or attitude measures (Schmeidler and McConnel, 1958) or by correlating ESP scores with relevant variables. Traditionally, when correlating ESP scores with other variables there has been no suggestion that overall ESP scores should show a significant deviation from MCE (e.g. Feather, 1967). Sargent(1979) has argued that it is by just such techniques that we can detect the ESP otherwise unobservable in chance scores.

2. An alternative argument is that one might be simply correlating random scores with some other measure and hence producing meaningless correlations. We could adopt the general guideline that the validity of any measure cannot be higher than its reliability. In this case the validity of the relationship between ESP scores and memory test scores cannot be expected to be higher than the reliability of the scores on which it depends. The memory scores are likely, although untested here, to be fairly reliable, but the same cannot be said of the ESP scores without testing. In this experiment we can test the reliability of the ESP measure in two ways.
(a) Since each S completed 2 ESP sessions we can calculate the test-retest reliability of the ESP scores. This gives $r=-0.49$ ($t=2.36$, 19df, $p=0.03$). This correlation is significant but is in the opposite direction to that expected. Ss tended to score in the opposite direction on the two occasions. The intertest reliability of the ESP score is therefore negligible.

(b) It could be argued that reliability could not be expected between ESP tests carried out so far apart in time. We can therefore test the reliability within one experiment by halving the data for each S and correlating the two halves. In the second session there were four runs of 25 trials each. For each S run scores on the first and third runs were compared with those on the second and fourth. This gives a correlation of $r=-0.17$ ($t=0.64$, 19df, $p= 0.53$) which is not significant. The correlation is small and negative and implies there is no internal consistency to the ESP scores in this case.

Of course in other experiments the reliability may be higher and certainly more Ss should be tested before reaching any conclusions but with no test-retest reliability or internal consistency I believe it must be unwarranted to place any reliance on a correlation between this measure and any other.

If we follow this second argument we may conclude that we should only look for correlations between ESP and other measures if we have either (a) evidence from overall scores that ESP occurred or (b) some indication of reliability in the data. This was not the case in this experiment. The following, similar, experiment used an improved method and more Ss.

EXPERIMENT 2

36 students in a parapsychology class completed 2 ESP tests and a memory test. The ESP tests each consisted of 4 runs of clairvoyance with individual target lists sealed in opaque envelopes. The memory test was similar to that used in the previous experiment except that only 3 minutes were allowed for learning and recall of the two sets of 25 words. Words in the first set were all of 6 letters or less and in the second set of more than 6 letters. In the final recognition test Ss were given 3 minutes in which to underline those words embedded in 100 words, which had appeared in the first set. The number of words from the first set recalled and recognised was recorded and the recall-recognition ratio calculated. Correlations between these memory scores and ESP scores were calculated as before.

Results

The mean ESP score was 41.6 which is not significantly different from MCE of 40.0 (t=1.98, 35df, p=0.06). For the memory tests means were 17.9 for recall, 20.6 for recognition, and 0.87 for recall-recognition ratio. When these are correlated with the ESP scores the following results are obtained:

Recall	r = 0.43	t = 2.78	35 df	p = 0.009
Recognition	r = 0.31	t = 1.90	35 df	p = 0.07
Ratio	r = 0.29	t = 1.77	35 df	p = 0.09

These results are different from those of the previous experiment. Memory scores are lower, presumably because less time was allowed for learning, and there is a significant positive correlation between ESP and recall scores, though not with the other measures.

Discussion

These results indicate a positive correlation between ESP and word recall ability, though not recognition. This outcome was not expected on the basis of any of the models. It implies that there may be some process involved in recall, but not in recognition, which is relevant to ESP. This is likely to be a process concerned with retrieval rather than storage. At first sight this might indicate support for a memory theory of ESP such as Carington's or Price's rather than Roll's memory theory of ESP. In this respect it is interesting that the positive correlation was obtained with the harder memory task, which conforms to the suggestion that the more difficult memory task more closely resembles ESP. This possibility was tested directly in experiment 6.

Before trying to interpret these findings further, though, some words of caution are needed. First, as previously, the ESP tests were conducted before the memory tests and Ss knew their ESP scores, if they could remember them. This problem was eliminated in later tests. Secondly we can check the reliability of the ESP measure used here by determining the correlation between the two sets of ESP scores. For these r=-0.09, 35df, t=0.55, p=0.59. This correlation is very small

and it must be concluded that the ESP scores used do not constitute a reliable measure. In further experiments different memory tasks were used.

EXPERIMENT 3

23 students in a parapsychology class completed 2 ESP tests each consisting of 200 trials of clairvoyance with ESP cards, and a test of incidental learning. Before a parapsychology lecture various objects were placed on two tables in the lecture theatre, words and numbers were written on the blackboard and pictures in different colours were drawn on the overhead projector. During a break these were removed and on returning Ss were given an answer sheet and asked to recall as many of the 20 items as they could. For each S the number recalled correctly was correlated with his ESP score.

Results

The mean ESP score was 41.1 which does not differ significantly from MCE of 40.0 ($t=1.09$, 22df, $P=0.29$). The memory test was effective in that most Ss recalled some items although none recalled all 20 items correctly. The mean number recalled was 7.8. When memory scores are correlated with ESP scores $r = -0.40$. This is not significant ($t=2.0$, 22df, $p=0.06$).

Discussion

No significant correlation was obtained between ESP scores and scores on an incidental learning test. However, the correlation obtained was in the opposite direction to those obtained for recall and recognition and this difference was pursued in later experiments.

There were again two major faults. In the ESP test all Ss were given one target order, and Ss knew their ESP scores before taking the memory test. These faults were eliminated in the next, similar, experiment.

EXPERIMENT 4

50 students in a parapsychology class completed 200 trials of clairvoyance, 100 trials prior to the memory test and 100 trials afterwards, with individual target lists sealed in opaque envelopes. The memory test was the incidental learning task used in the previous experiment.

Results

The mean ESP score was 41.64 which is significantly above the MCE of 40.0 ($t=2.35$, 49df, $p=0.023$). Memory scores varied from 2 to 17 items recalled with a mean of 8.26. The correlation between these scores and the ESP score for each S gives $r=-0.19$ ($t=1.34$, 49df, $p=0.18$). This correlation is not significant but it is negative, like that found in the previous experiment, which may indicate some weak effect.

It has previously been suggested that the reliability of the ESP measure should be tested. In this experiment the combined ESP scores are significantly above MCE. However, the intertest reliability is again low with $r=0.10$ ($t=0.70$, 49df, $p=0.49$). This means that although the overall ESP scores are significant the measure is not reliable and any correlation based on it may not be reliable either.

In previous experiments it was mentioned that Ss' prior knowledge of their scores might produce spurious correlations with memory scores. In this experiment half the test was conducted before the memory test and half afterwards. This makes it possible to use the ESP scores separately in correlations to see whether there is any difference. The results are as follows:

Memory - ESP(1)	$r = -0.16$	$t = 1.09$	49 df	$p = 0.28$
Memory - ESP(2)	$r = -0.11$	$t = 0.77$	49 df	$p = 0.45$

There is little difference between these two and it can be concluded that the Ss' knowledge of their ESP scores did not significantly affect their memory scores. This problem therefore does not seem important although it would be preferable to avoid it altogether if possible. In the next experiment one ESP test was carried out before

the memory test and one in the same session. The two could be compared to see whether there is any difference if the two tests are taken at the same time.

EXPERIMENT 5

20 students in a parapsychology class completed in one session 100 trials of clairvoyance with individual target orders sealed in opaque envelopes, and in a second session a memory test followed by a further 100 trials of clairvoyance. The procedure for the memory task was identical to that used in the last experiment, that is, an incidental learning task.

Results

The mean ESP score for the first test was 19.70, which is not significantly different from MCE of 20.0 (t=0.36, 19df, p=0.73). For the second ESP test the mean was also 19.70 (t=0.34, 19df, p=0.74). For the combined results the mean is 39.4 where MCE is 40.0 (t= 0.53 19df p= 0.60). The mean number of items correctly recalled in the incidental learning test was 12.4. Various correlations of interest are listed below.

Memory - ESP (combined)	r=-0.22	t=0.96	19df	p=0.35
Memory - ESP (1)	r=-0.18	t=0.77	19df	p=0.45
Memory - ESP (2)	r=-0.18	t=0.50	19df	p=0.62
ESP (1) - ESP (2)	r=-0.19	t=0.50	19df	p=0.62

The correlation between total ESP scores and memory scores is again negative but is very small. Two other points are of interest.

1. The intertest reliability of the ESP scores is low and the correlation between the two negative. To this extent the ESP measure is unreliable.
2. It was suggested that a test of ESP made at the same time as the memory test might produce a higher correlation. However, the results of the test given immediately after the memory test, and before the results of this were known, shows no difference from the earlier test.

The means for the two tests are the same and the correlations with memory scores similar and both insignificant. Clearly, carrying out the test in the same session makes no difference to the results.

Discussion

The results of three similar studies can now be summarised. The correlations between ESP and incidental learning scores are shown below.

1.	Experiment 3	r=-0.40	t=2.00	22df	p=0.06
2.	Experiment 4	r=-0.19	t=1.34	49df	p=0.18
3.	Experiment 5	r=-0.22	t=0.96	19df	p=0.35

All the correlations are negative but none is significant. In addition it should be noted that experiment 3 used a faulty ESP test procedure. The timing of the ESP tests was varied but the intertest reliability of the ESP tests was uniformly low and it appeared to make no difference when the tests were carried out. However, since all the correlations with incidental learning were negative it was thought this might indicate some weak effect and so the final experiment also used an incidental learning procedure but of a different kind.

MAIN STUDY EXPERIMENT 6

In previous experiments a positive correlation between recall scores and ESP and a small negative correlation between incidental learning and ESP scores were obtained. Although all the correlations were small, if we assume, for the moment, a real effect, then several possible differences might be responsible. These include

1. Attentional differences.
2. The different extent to which the information was processed.
3. The difficulty of the memory task.

Undoubtedly there are other possibilities which could be tested but it was thought premature to try to test any one hypothesis in detail. Instead the following experiment was designed as a general test for all the above hypotheses. Ss were given an ESP test with words as targets and afterwards were unexpectedly asked to recall both those

words which they had chosen as targets in the ESP test and those they had not chosen. The two scores were then correlated separately with ESP scores. The chosen words might be expected to be better attended to, better processed, and easier to recall than the non-chosen words. We should therefore expect a difference in the correlation between the two types of memory test and the ESP scores. That is, if any of the above hypotheses is correct, the correlation between the number of non-chosen words recalled and ESP should be higher than that between the chosen words recalled and ESP. If such a difference is not found these hypotheses can be rejected, but if it is found then further experiments should be designed to explore them further. Since this task is again one of incidental learning the minimum prediction is that both correlations should be negative.

Method

Ss were 32 students in a parapsychology class. Targets were 50 common four-letter words arranged in 10 groups of five. On each trial Ss were asked to choose one of five words. The target lists of 10 words each were prepared by computer and sealed in opaque, numbered envelopes by an assistant who took no further part in the experiment. Ss were given a list of 10 groups of five words each and an envelope containing their target list. When they had all finished they gave in their sheets and were given a blank piece of paper and asked to remember as many of the words as they could, including those they had chosen as targets and as many others as possible. When they had finished, the test sheets were returned to other students to be marked. All were rechecked later. For each S the ESP hits, and the number of words, both chosen and not chosen, was recorded.

Results

The mean ESP score was 1.94 which is slightly, but not significantly, below MCE of 2.0 ($t=0.32$, $p=0.25$). Memory scores for the chosen words averaged 7.5 (75%, sd=2.89) and for the non-chosen words 6.2 (15.6%, sd=4.56). This large difference indicates that, as expected, the chosen words were more readily recalled and had presumably been better attended to. The two tasks therefore differed in difficulty as intended. The mean for overall memory scores was 13.72 (sd=5.60). For each S the two memory scores were correlated with

the ESP test score and the following correlations were obtained:

ESP - chosen words recalled	r=0
ESP - non-chosen words recalled	r=0.32 (z=1.80, 31df, p=0.07)
ESP - total words recalled	r=0.26 (z=1.46, 31df, p=0.14)

None of these correlations is significant.

Discussion

Small correlations were again obtained and the difference between the two correlations was in the direction opposite to that expected. That is, negative correlations were predicted with the higher correlation between ESP and non-chosen words. Instead there was no correlation between the number of chosen words recalled and ESP and a positive, though small, correlation between ESP and the number of non-chosen words recalled. The main hypothesis of this experiment was therefore not confirmed.

Additionally, from previous results, we can make the simple prediction that since this task was one of incidental learning the correlations would be expected to be negative. None of them was and the overall memory score gives a weak positive correlation with ESP scores. This makes it seem more likely that previous weak negative correlations were spurious, or at least, if there was an effect operating it is either specific to the previous task used and not generalisable to other incidental learning tasks or it is sporadic and not repeatable.

CONCLUSION

The results of 6 experiments have been reported here. On the basis of the preliminary experiments several hypotheses were made and tested in a final experiment but were not confirmed. The most reasonable conclusion seems to be that the weak correlations obtained were due to chance.

SUMMARY AND DISCUSSION OF RESULTS

The results of 6 experiments investigating correlations between ESP and memory scores have been reported. Table 1 is an attempt to summarise these results. Several conclusions are possible.

1. The most conservative argument is that correlations can only be meaningful if the measures correlated are reliable and consistent. It has been shown that the ESP measures used here are neither consistent nor reliable. The corresponding correlations might therefore be meaningless.

2. One may accept the unreliability of the ESP measure but allow only those correlation coefficients with an associated probability of less than 0.05 to be taken as indicative of a real effect. Here we have one, in experiment 2, which shows a positive correlation between recall and ESP scores. This could lend support to a memory model of ESP and lead to speculation on the relationship between this particular task and ESP.

TABLE 1
Results of 6 experiments correlating ESP and memory

Experiment	Overall signif	Correlation of ESP with Recall	Recognition	Ratio	Memory task Inc le	1 or 2
1	no	0.09	0.002	0.21	-	secondary
2	no	0.43	0.31	0.29	-	secondary
3	no	-	-	-	-0.40	secondary
4	$p<0.05$	-	-	-	-0.19	secondary
5	no	-	-	-	-0.22	secondary
6	no	-	-	-	0.26	both

However, caution is needed here, firstly because the effect is obviously not stable, since there was no sign of it in experiment 1,

which used the same task, and secondly because this is only one significant finding among so many. In 6 experiments overall ESP scores were significant in only one, experiment 4, and of 10 independent correlation coefficients reported as primary analyses only one is significant, in experiment 2. Furthermore, if all analyses are considered (though not all are independent) in a total of 34 significance tests 2 were significant at $<.05$. This does not seem good grounds for rejecting the null hypothesis that chance alone accounted for all the results.

3. It can finally be argued that the effects of ESP are extremely weak and often obscured by counteracting effects but nonetheless the directions of the correlations can give us valid information. Following this argument the data here indicate two lines of approach. The first six experiments all conformed to the pattern that deliberate learning correlated positively with ESP while incidental learning correlated negatively. Simple hypotheses were developed from this finding and experiment 6 was designed to test them by comparing two different memory tasks and an ESP test. The results were not as predicted and a weak positive correlation between ESP and incidental learning was found.

Another possible way of interpreting these results was suggested recently, and after these experiments were completed, by Irwin (1979). He suggested that measures of secondary memory may correlate positively with ESP, and measures of primary memory negatively. In Table 1, the type of memory task is indicated, and it appears that Irwin's suggested framework does not fit these data.

Even if one ignores unreliability and levels of significance and looks only at direction of correlations there does not seem to be any meaningful pattern in these data. But this is perhaps not surprising. The attempt to correlate ESP with other measures may, after all, be futile. If no ESP occurs it is clearly impossible to test hypotheses about its nature and even if it does occur the predictions must be clearly formulated if any conclusions are to be possible.

In conclusion it seems that the results of 6 experiments on the correlation between ESP and memory provide no evidence for any such correlation and lead to no hypotheses for future research. I would go so far as to say that the data presented here appear to describe the purely random results of pairing chance scores with memory test scores. In the absence of any ESP it is clearly impossible to test hypotheses about the nature of the process. To do this it will be

necessary to find a source of ESP first.

ABSTRACT

It has been claimed that several theories of ESP predict a positive correlation between ESP and memory. The basis of these predictions is discussed and the existing evidence for such a correlation assessed. Ways of making sense of the conflicting findings are considered and some further predictions drawn. 6 experiments are reported. In two preliminary experiments, recall and recognition scores were correlated with ESP scores and one positive correlation with recall obtained. In three further studies negative, but non-significant, correlations between ESP and incidental learning were found. From these findings it was predicted that a negative correlation would be found between ESP and incidental learning scores with the higher correlation with the harder task. These predictions were tested in the main expeiment but were not confirmed. The ESP measures used were found to be inconsistent and unreliable and it was concluded that hypotheses about the nature of ESP could not be tested in the absence of any ESP.

REFERENCES

Bergson, H. 'Matter and Memory', Trans. Paul and Palmer. Swan Sonnenshein, London, 1911.

Blackmore, S.J. 'Extrasensory perception as a cognitive process', Unpublished PhD thesis, University of Surrey, 1980.

Carington, W. 'Telepathy :An outline of its facts, theory and implications', Methuen, London, 1945.

Feather, S.R.'A comparison of performance in memory and ESP tests', in 'Parapsychology from Duke to FRNM', Rhine, J.B. et al., Parapsychology Press, Durham, N.C., 1965, 78-81.

Feather, S.R. 'A quantitative comparison of memory and psi', Journal of Parapsychology, 1967, 93-98.

Irwin, H.J. 'On directional inconsistency in the correlation between ESP and memory', Journal of Parapsychology, 1979, 31-39.

Kanthamani, H. and Rao, H.H. 'The role of associative strength in memory-ESP interaction', Journal of Parapsychology, 1975, 1-11.

Marshall, N. 'ESP and memory: A physical theory', British Journal for the Philosophy of Science, 1960, 265-286.

Parker, K. 'A study of immediate memory and ESP', in Morris, J.D. et al (eds) Research in Parapsychology 1975, Scarecrow Press, Metuchen, N.J. 1976, 130-134.

Peterson, C.J. 'Relationships among attitude, memory, and extrasensory perception', Journal of Parapsychology, 1972, 94 (abstract).

Price, H.H. 'Haunting and the "Psychic Ether" hypothesis', Proceedings of the Society for Psychical Research, 1939, 307-343.

Rao, K.R. 'Further studies of memory and ESP', Journal of Parapsychology, 1978, 167-178.

Rao, K.R., Morrison, M. and Davis, J.W. 'Paired associates recall and ESP: a study of memory and psi missing', Journal of Parapsychology, 1977, 165-189.

Roll, W.G. 'ESP and memory', International Journal of Neuropsychiatry,1966, 505-512.

Stanford, R.G. 'Extrasensory effects upon "memory"', Journal of the American Society for Psychical Research, 1970, 161-186.

PROJECT CHICKEN LITTLE: A PRECOGNITION EXPERIMENT INVOLVING THE SKYLAB SPACE STATION

William Braud
Mind Science Foundation

> One day Chicken Little strutted through the woods on her way to market. A tiny acorn fell from a tall tree and hit the top of her little head.
> "My goodness! The sky is falling!" said Chicken Little. "I must rush to tell the king!".
> (from a popular children's story)

Spontaneous precognitive experiences or premonitions very frequently provide information about future events of great import or significance to the experient. Frequently, the precognized event is tragic in nature, involving a disaster, a death, an accident, an injury, or an illness. In controlled experimental studies of precognition, however, the to-be-precognized or target event is an essential trivial one, one without special meaning to the subject. This difference in the degree of importance (or interest) of the 'target' event no doubt accounts for much of the difference in accuracy and richness between precognitive experiences in everyday life situations and those contrived in the laboratory.

Ethical and practical constraints prevent the use, in the laboratory, of really potent and meaningful target events. Even in the

absence of these constraints, the presence of great individual differences among subjects in degree of preference for or aversion towards possible powerful target events makes the selection of useful targets quite difficult. Occasionally, however, Nature provides a solution to our dilemma in the form of a 'spontaneous' event which is rich in meaning, interest and importance, but which nonetheless may provide data amenable to quantitative analysis. Such an event was the recent fall to earth of the American SKYLAB space station.

SKYLAB was originally expected to remain aloft until 1983. However, increased atmospheric drag on the space station, along with progressive deterioration of its control systems, caused SKYLAB to fall to earth much earlier than originally predicted. At the time the experiment to be reported here was initiated, the reentry of SKYLAB was an essentially unpredictable (quasi-random) event. Official nonpsychic predictions (at that time) of the time of reentry ranged from April 1979 to early 1980. Although it was known that reentry would take place somewhere within the 50 N to 50 S latitude region covered by SKYLAB's orbit, the exact fall down site was not known because of uncertainties of the upper atmosphere and of SKYLAB's drag characteristics. The atmospheric uncertainty was contributed by unpredictable variations in the sunspot activity. It was estimated that the time of reentry might be predictable perhaps 12 hours in advance from early signs of tumbling, but the location might be uncertain until there was only an hour or less remaining.

At the time of the experiment, SKYLAB's reentry provided an event the time and place of occurrence of which could not be predicted by rational inference. SKYLAB's reentry also provided a very meaningful target event. First, the falling to earth of a large, manmade space station was a rare and unusual occurrence, one which generated a great deal of public interest and media coverage. Second, although NASA suggested that the likelihood of damage or injury from descending bits of SKYLAB was less than the chance of being hit by a meteorite, the possibility of some destruction could not be ruled out, particularly if SKYLAB descended over a highly populated area.

The existence of such a potentially important but unpredictable event suggested an interesting experiment. Could the then-unknown location and time of SKYLAB's impact be predicted psychically? To test this hypothesis, the Mind Science Foundation initiated 'Project Chicken Little', a precognition experiment involving the SKYLAB space station. The purpose of the project was twofold: (a) to determine whether the time and site of SKYLAB's impact could be accurately

precognized, and (b) to assess the relative accuracy of predictions for various prespecified subgroups of participants.

METHOD

On February 1, 1979 a letter and response sheet were mailed to potential participants. The letter (see Appendix A) described the project in detail. The respondent was asked to indicate his or her precognitive guess of the major impact site on a map (see Appendix B) of the earth's surface between 50 N and 50 S latitude. The respondent was also asked to predict the precise time of impact. Responses were to be mailed back to Mind Science Foundation for evaluation. The deadline for returns was set for May 1, 1979.

The potential participants were 150 parapsychologists, 'psychics', and other persons who were personally known by the author and judged to have an interest in participating in this project. His original intention was to compare the relative accuracy of three subgroups: parapsychologists vs 'psychics' vs other selected participants. If large enough samples responded, a comparison was to be made of 'psi-conducive' experimenters (i.e., those whose experiments frequently yield evidence of psi) and 'psi-antagonistic' experimenters (i.e., those whose experiments customarily yield no evidence of psi). However, certain participants, duplicated and distributed letters and response sheets to still other persons, and therefore a considerable number of participants beyond the original 150 became involved. At the time of the May 1st deadline, 200 responses had been received. Following the May 1st deadline, response sheets continued to arrive, most of these from distant lands, especially Brazil, Japan, and the United Kingdom. Since there was such a large number of these 'late' responses (there were 273 late responses from Brazil alone), it seemed wasteful to discard them; therefore, the deadline was extended to include all responses postmarked on or before July 10, one day before SKYLAB actually fell to the earth. A total of 512 responses were received before this new deadline. Since the actual impact site was still unknown even a few hours before falldown, the spatial responses of all 512 participants are 'safe' (i.e., not contaminated by possible knowledge of nonpsychic impact estimates which may have been obtained through media announcements). However, after the original May 1st deadline, NASA and NORAD began to make and continually update estimates of the most likely time 'window' for falldown. Since some participants may have had access to this late information which

effectively narrowed the possible time frame for the event (compared with the original time range given in the initial letter), a conservative strategy was adopted of excluding all temporal responses received after the original May 1st deadline. Thus, the final data sample included 512 impact site estimates (200 received before the first deadline and 312 received before the second deadline), but only 200 impact time estimates (the 200 received before the first deadline).

Before the data were observed, the participants were divided into five groups. The first group (n=35) consisted of all parapsychologists who responded to the author's original invitation to participate; this was one of the originally specified groups. The second group (n=327) consisted of 'trained' volunteers. Nine of these were well-known 'psychics' with long histories of impressive psychic functioning; this was one of the originally specified subgroups. The remainder were persons who had 'graduated' from a program (Silva Mind Control) claimed to facilitate psi functioning. The original letter and map were reproduced in various Mind Control newsletters, and graduates who wished to participate were requested to mail their responses directly to the Foundation. Three hundred eighteen Mind Control graduates participated; 26 of these were located in the United States and 292 were located in Brazil. The third group (n=25) consisted of participants specially selected by the author. These participants were local San Antonio residents who had performed well in prior Mind Science Foundation experiments and who had maintained a serious interest in the Foundation's work. This was the third of the originally specified subgroups. The fourth group (n=93) consisted of members of groups having a special interest in psi. In particular, this group consisted of 22 members of the Japanese Society for Parapsychology, 18 persons enrolled in a survey course in parapsychology in Connecticut, 11 persons enrolled in a parapsychology course in Chicago, 18 members of the Texas Parapsychological Association (a local organization of interested nonprofessionals), and 24 respondents from the United Kingdon who were members of various psychic study groups. The fifth group (n=32) consisted of miscellaneous unselected subjects who simply responded after learning of the experiment from friends and acquaintances.

The map of the earth's surface (see Appendix B) is divided into twenty-four 15 degree longitudinal segments. Each participant's impact site estimate could be assigned to one of these 24 segments. Since each SKYLAB orbit carried it over all 24 segments, one might assume that the probability of an estimate segment corresponding to the

segment in which the major impact actually occurred was 1/24 or .0417. Thus, one could count the number of predictions falling within the actual impact site segment and compare this obtained frequency with the frequency expected on the basis of chance alone. A frequency polygon would indicate the distribution of the 512 predictions. If one assumes equiprobability of the 24 segments as potential impact sites, the foregoing method could be used to infer the presence of a psi effect. There is no correspondingly simple method by which the time estimates may be evaluated for psi; and, in fact, the time estimates were used only in an analysis of the relative accuracy of the five subgroups.

A major aim of this study was a comparison of the five prespecified subgroups in terms of the relative accuracy of their spatial and temporal predictions. Since each participant indicated his or her spatial prediction by means of an 'x' on the map, spatial accuracy could be readily assessed by simply measuring (to the nearest millimeter) the distance between the 'x' and a point indicating the actual impact site. If an 'x' was more than halfway around the globe from the impact site, the map was 'folded over', and the shorter distance was used as the measurement. This was accomplished by aligning the right edge of one map with the left edge of another and measuring the shortest distance from the 'x' to one of the two representations of the actual impact point. The accuracy of a temporal prediction was assessed by counting the number of days between the estimated date of impact and the actual impact date. This computation was aided by a special software program run on a Texas Instruments Programmable 59 calculator.

RESULTS

Shortly after the May 1st deadline, the 200 spatial and temporal predictions which had been received at that time were tabulated and graphed. Since it was important to make the predictions known before the predicted event occurred, a preliminary report of the results of the experiment was prepared, and on May 20, 1979, copies of this report were mailed to the Central Premonitions Registry, the recently formed Alpha Premonitions Bureau (U.K.), the Society for Psychical Research, the American Society for Psychical Research, and the Committee for the Scientific Evaluation of Claims of the Paranormal. Since, of course SKYLAB had not yet fallen to earth, the only findings that could be reported at that time were the frequency distributions

for the spatial and temporal predictions. It might be useful to summarize these here.

A frequency polygon indicating the total number of predictions for each of the 24 longitude segments is given in Figure 1. As can be seen, the impact site predictions are characterized by tremendous variability. The appropriate measure of central tendency for these nominal data is the mode. The modal segment, containing 18 (9%) of the 200 predictions, is the segment bounded by 60 W longitude and 75 W longitude. This figure contrasts with the 8.33 (4.17%) predictions expected on the basis of equal probability of the 24 segments. A binomial test (corrected for continuity) applied to the frequencies for this segment yields a $z=3.24$, $p=.0012$, two-tailed. Were this the only information upon which we could rely at that time, we would predict that SKYLAB would fall to earth somewhere in this particular segment.

The predictions for time of impact are given in Figure 2. Again, these estimates are characterized by considerable variability. Estimates of the month of SKYLAB's descent ranged from February, 1979, through November, 1981. The three measures of central tendency for the 195 predictons (five persons neglected to make temporal predictions along with their spatial predictions) agree rather well. The modal month is August, 1979, with 20 (10.26%) of the score. October, 1979, is the month containing both the mean and median scores. Thus, having only this information, we would predict that SKYLAB would descend in either August or October of 1979.

By July 10, an additional 312 responses had been received, changing the frequency distribution of spatial predictions to that shown in Figure 3.

The modal segment now became the one bounded by 0 and 15 W longitude, containing 44 (8.59%) of the 512 predictions. This figure contrasts with the 21.33 (4.17%) predictions expected on the basis of equal probability of the 24 segments. A binomial test (corrected for continuity) applied to the frequencies of this segment yields a $z=4.90$, $p=9.6E-7$, two-tailed.

The results for the temporal predictions for the 512 returns are identical to those for the original 200 returns, since the 312 'late' temporal estimates were excluded from the analyses. The reason for this exclusion has already been discussed.

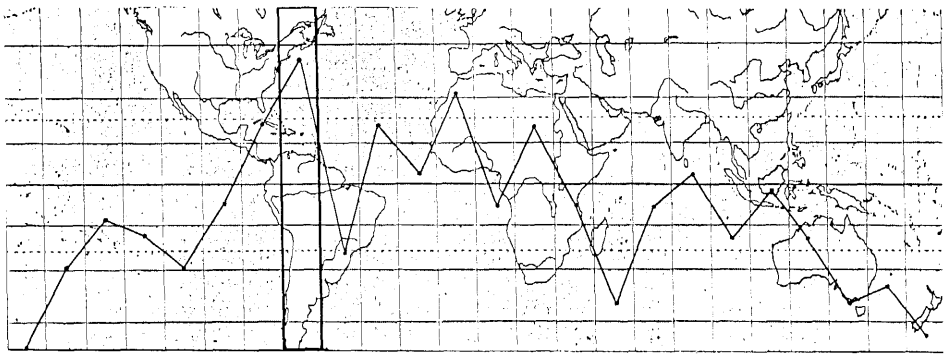


FIGURE 1
Frequency polygon indicating the number of impact site predictions made for each of 24 longitude segments by the 200 pre-May first deadline participants

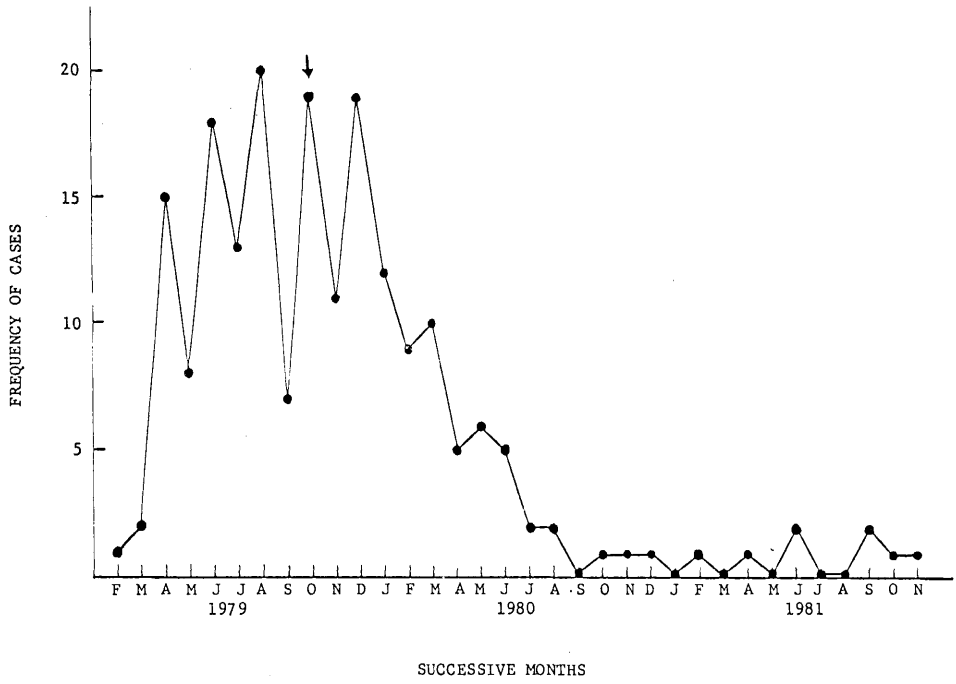


FIGURE 2
Frequency polygon indicating number of impact time predictions for each month from February 1979 through November 1981. Arrow indicates month with both median and mean number of predictions

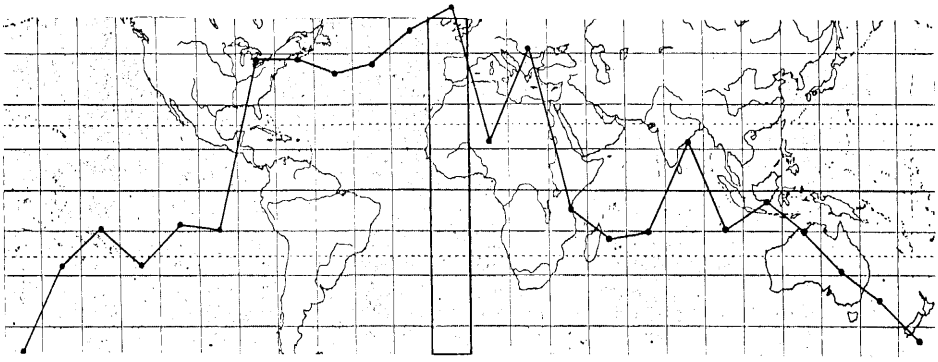


FIGURE 3
Frequency polygon indicating the number of impact site predictions made for each of 24 longitude segments by all 512 pre-July 10th deadline participants

SKYLAB actually fell to earth, not within either of the two modal segments discussed above, nor within the months of August or October of 1979, but rather near Kalgoorlie, Australia, on July 11,1979. Kalgoorlie lies within the segment bounded by 120 E longitude and 135 E longitude. Thus, the departures from chance expectation of the frequencies within the modal segments and months mentioned above reflected something other than accurate psychic awareness of the impact location and time. In fact, the actual impact segment contained 15 (2.92%) of the 512 spatial predictions, a frequency which does not differ significantly from the 21.33 (4.17%) predictions expected by chance (z=-1.29, p=.197, two-tailed, corrected for continuity).

TABLE 1
Separation (in mm.) of predicted and actual reentry sites for five groups of participants

Group	N	Mean Distance
Parapsychologists	35	87.03
'Trained' Ss	327	85.86
Selected Ss	25	86.36
Interest groups	93	79.57
Miscellaneous unselected Ss	32	81.62

The relative accuracy of the spatial predictions (i.e., distance measurements) of the five subgroups is indicated in Table 1. A simple, one-way analysis of variance indicates that the difference among these five groups are not significant (F=0.82, df=4/507, p=n.s.).

The relative accuracy of the temporal predictions (i.e., number of days between actual and predicted reentry date) for the five subgroups is indicated in Table 2. Only the 200 predictions received before May 1st are included here (actually, 195 scores are included, since five persons neglected to make temporal predictions). Although the differences among the means for this time measure are greater than

those for the spatial measure, a simple, one-way analysis of variance indicates that these differences are not significant ($F=2.21$, $df=4/190$, p=n.s.).

TABLE 2
Separation (in days) of predicted and actual reentry times for five groups of participants

Group	N	Mean time (days)
Parapsychologists	26	138.77
'Trained' Ss	46	137.39
Selected Ss	24	101.29
Interest groups	70	181.36
Miscellaneous unselected Ss	29	216.41

DISCUSSION

The participants in this study did not accurately predict the actual impact site and time of impact of the SKYLAB space station. The deviations from chance of the frequency measures plotted in Figures 1 and 3 appear to reflect something other than precognitive accuracy, perhaps a response bias to select sites near the center of a display or a tendency to avoid the extremes of a display (a map, in this case).

There exists a possibility, however, that the two peaks of Figures 1 and 3 may reflect psychic predictions after all. Twelve hours before falldown, one NASA-NORAD projection of the likely time and place of SKYLAB's impact seemed to threaten populated areas of southern Canada or Maine. At that point, a command was given to 'tumble' the space station -- "to turn off control gyros to create low altitude instability that would prolong reentry and provide a more predictable trajectory" (Williams et al., 1979, p. 33). The peak in Figure 1 may

reflect a prediction of where SKYLAB may have impacted had this last minute correction not been made. This admittedly quite speculative interpretation carries with it an interesting suggestion: that participants may have been, in some sense, working forward in time, extrapolating from past and present information, rather than making their predictions solely on the basis of some direct apprehension of the future event itself. This notion may have application in the context of the intervention problem in precognition.

According to a map published in Newsweek (Williams et al., 1979, p. 31), SKYLAB actually began to break up approximately over Ascension Island. The peak of Figure 3 may reflect psychic awareness of this event. Again, this suggestion is admittedly extremely speculative.

Although two potentially important events during SKYLAB's final orbit may have been psychically detected, there was no indication in the available data of awareness of the final impact site itself. The absence of any injuries due to impact and the relatively long time span between the time participants made their predictions and the time of the predicted event may have contributed in some way to the absence of an obvious psi effect in these data. Perhaps the final event was not as meaningful as originally assumed.

An additional factor which may have obscured a true precognition effect in these data is a possible psychokinetic effect upon SKYLAB. A person predicting a particular impact site may have unintentionally exerted a psychokinetic influence upon the sunspot-atmosphere-spacecraft system in such a manner as to increase the likelihood of actual impact at that site. Since many persons were involved in the study, generating different predictions, and hence different 'PK target sites', the various PK influences may have somehow cancelled each other. The author realizes that this is a rather extreme hypothesis. However, in opposition to such an hypothesis one can only put forward an assumption regarding the limits of psychokinesis. We have no empirical knowledge of those limits.

The author's expectation was that the five prespecified groups of participants would differ in the degree of accuracy of their predictions. The expectation was that parapsychologists would be most accurate, followed by 'trained' subjects, selected subjects, interest groups, and miscellaneous unselected subjects, in that order. Inspection of the means of Tables 1 and 2 and the ANOVA results indicates that this expectation was not confirmed.

A final comment might be made regarding a possible 'stacking effect' in an experiment such as this one. The fact that the predictions were made independently by many persons at very different times and at very different locations minimizes the likelihood of a certain kind of stacking effect. The fact that there was only one target event raises the question of how to distinguish a true precognition effect from a response bias effect. Perhaps a response bias effect could be minimized (in future studies of this type) by presenting the response display (the map, in this case) in slightly different forms to different respondents or by asking participants to make their responses in slightly different ways. This strategy might result in dilution of response bias while maintaining any true precognitive effect. However, since such an experiment necessarily involves a single target event, inherent uncertainty remains, as in nearly all 'spontaneous' psi occurrences. Perhaps such experiments should be viewed as 'quasi-experimental' or 'controlled spontaneous' events, and investigated primarily for their heuristic value whenever exceptionally interesting, incompletely predictable 'accidents' present themselves.

SUMMARY

Five hundred twelve persons participated in an experiment in which they attempted to predict the exact time and place of occurrence of a single quasi-random event, the fall to earth of the American SKYLAB space station. These participants were not able to accurately predict the actual impact site and time. Five prespecified subgroups of participants (parapsychologists, 'trained' subjects, selected subjects, interest groups, and miscellaneous unselected subjects) did not differ in relative accuracy of spatial or temporal predictions.

There were present, however, certain indications of possible psychic awareness of other orbital events, which are relevant to the so-called 'intervention' problem in precognition. These indications were discussed, along with certain methodological problems (stacking effect and response bias).

ACKNOWLEDGMENT

I am grateful to Gary Davis for his assistance during several stages

of this experiment.

REFERENCE

Williams, D.A., Hager, M., Alexander, L. & Robinson, C. 'SKYLAB's fiery plunge', Newsweek, 1979, July 23, 31-33.

Willian Braud
Mind Science Foundation
102 West Rector, Suite 215
San Antonio, Texas 78216
U.S.A.

BOARD OF TRUSTEES
WILLIAM T. RHAME
CHAIRMAN
GEORGE W. CHURCH, JR.
MARTIN GOLAND
FRANK HARRISON
ROBERT P. MCCULLOCH
BETTY SLICK MOORMAN
HARRISON A. PRICE
TOM SLICK
JERI WALSH
DUNCAN WIMPRESS
C. V. WOOD, JR.

MIND SCIENCE FOUNDATION
102 W. RECTOR STREET
SAN ANTONIO, TEXAS 78216
(512) 349-5127

EXECUTIVE COMMITTEE
JERI WALSH
CHAIRMAN
WILLIAM T. RHAME
C. V. WOOD, JR.

OFFICERS
TOM SLICK
PRESIDENT
C. V. WOOD, JR.
VICE PRESIDENT
JOHN O. R. JOHNSON
TREASURER
JOANNE WOOD
SECRETARY

1 February 1979

TO: All potential participants

SUBJECT: PROJECT CHICKEN LITTLE
A Precognition Experiment Involving the SKYLAB Space Station

Dear Friend:

On February 8, 1974, the final U.S. astronaut crew left Skylab, our huge orbiting space station. Skylab was expected to remain aloft until 1983, leaving plenty of time for our space shuttles to reach it for various purposes. However, increased drag on the space station, along with progressive deterioration of its control systems, indicate that Skylab will fall to earth much earlier than originally predicted. The most recent data indicate that Skylab will make its fiery descent through our atmosphere and fall to earth in late 1979 or early 1980. If sunspot activity should increase the density of the upper atmosphere, the increased drag on the orbiting workshop could result in an even earlier reentry time, perhaps even as early as April or May of 1979.

Skylab weighs about 175,000 pounds. Although most of Skylab will burn up in the atmosphere, some substantial chunks are likely to reach the earth's surface. Analysis indicates that 40,000 to 50,000 pounds of debris could survive reentry. The largest pieces that could hit the earth intact are the airlock shroud (at about 5,000 pounds) and Skylab's lead film vault (at about 4,000 pounds). Such large pieces could hit the earth with speeds as high as 200 to 300 feet per second.

On December 15, 1978, NASA recommended the abandonment of all further attempts to control Skylab. According to NASA's evaluation, more debris would survive reentry if Skylab were broken up in orbit than if the spacecraft were allowed to reenter as a single piece, with breakup occurring naturally.

Now, the interesting thing about Skylab's fall to earth is that the exact time and place of falldown are not known. We do know that reentry will take place within the 50°N to 50°S latitude band covered by Skylab's orbit. But this region (indicated on the enclosed map) includes a great deal of land: parts of North America (including most of the United States), Europe and Asia, most of South America, and all of Africa and Australia. Three-fourths of the area is ocean. We do not know exactly where the largest fragment or collections of fragments will fall. The exact site is unpredictable because there are uncertainties about the upper atmosphere and because Skylab's drag characteristics are poorly understood. A NASA official estimates that the time of reentry may be predictable perhaps 12 hours in advance from early signs of tumbling, but that the location may be uncertain until there is only an hour or less remaining.

Thus, the reentry of Skylab provides us with an essentially unpredictable (quasi-random) event which could be of considerable importance to a number of people.

APPENDIX A

Although NASA suggests that the likelihood of damage or injury from descending bits of Skylab is less than the chance of being hit by a meteorite, the possibility of some destruction cannot be ruled out. The existence of such a potentially important but unpredictable event suggests a very interesting experiment. Can we predict psychically the presently unknown impact location and impact time of Skylab? The purpose of this letter is to provide you with a unique opportunity to demonstrate precognition.

If you wish to participate in this experiment, here's what to do:

1. Examine the enclosed map, which indicates the 50°N to 50°S area of possible reentry; don't pay much attention to the grid lines (this happened to be the only world map that was readily available to us).

2. Make your prediction about the point of greatest impact by marking the map with an "X" at the predicted location. We realize that Skylab debris will not fall at only one place; debris will be spread over a large area along the reentry track. We are asking you to predict the location of the greatest amount of debris. Make this prediction psychically, on the basis of a hunch, impression, intuition or feeling. Do not try to figure out the "target" location logically. Do not second-guess yourself. Mark the map at one and only one place according to your first psychic impression.

3. In the blank space below the map, indicate your psychic guess of the exact time and date of the greatest impact.

4. Fill in the blanks in the space above the map.

5. Fold the map page, place it in an envelope, and mail it to us at this address:

PROJECT CHICKEN LITTLE
Mind Science Foundation
102 W. Rector Suite 215
San Antonio, Texas 78216

It is important that you mail your prediction to us as soon as possible, so that we can evaluate results well in advance of reentry. You will be notified about the accuracy of your prediction and about the outcome of the experiment as a whole. Thank you very much for your cooperation in this unique parapsychological project.

Sincerely yours,

William Braud

William G. Braud, Ph.D.
Research Associate

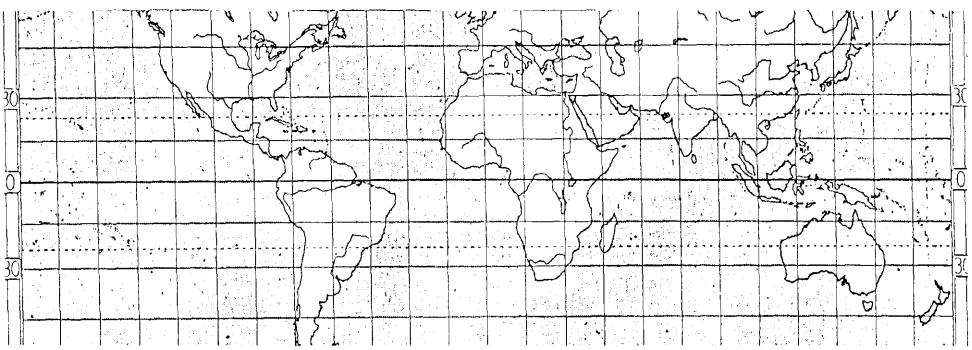
Enclosures: map page, description of Mind Science Foundation and its personnel

APPENDIX A

NAME _____ AGE _____ SEX _____

ADDRESS _____

DATE ON WHICH YOU ARE MAKING YOUR PREDICTIONS (I.E., TODAY'S DATE) _____



On the map above, please place an "X" to indicate your psychic guess of the location of greatest impact of the Skylab space station debris.

In the blanks to the right, indicate your psychic guess of the exact time and date of impact of Skylab space station debris

TIME _____ AM / PM ____ day ____ month ____ year

APPENDIX B

PK EFFECTS ON PRE-RECORDED GROUP BEHAVIOR OF LIVING SYSTEMS

Elmar R. Gruber

Institut fur Grenzgebiete der Psychologie und Psychohygiene
Freiburg, W-Germany

INTRODUCTION

The purpose of the experiments reported in this paper was to find out if a PK influence can be found on the pre-recorded behavior of a group of living systems. The original version of this paper, presented at the 1979 P.A. Convention in Moraga, California, was interpreted in the framework of Rex Stanford's model of "conformance behavior" (Stanford, 1977a; 1977b; 1978). Based on discussions with Stanford (1979) I decided to treat the experiments in this paper as PK experiments, though they might reflect conformance behavior.

The experiments are further studies concerning the question whether there exists a PK influence on the pre-recorded behavior of living systems as has been shown for animals and humans in recent studies (cf. Braud 1979a; 1979b; Gruber 1979a; 1979b). In these studies PK effects on the pre-recorded behavior of living systems could be found to some extent, but not in all experiments.

In the present study six experiments were performed with a total of 21 subjects. Three experiments involved as "biological REG" persons entering a supermarket on Saturday mornings. The remaining three experiments involved as "REG" cars passing through a small and short tunnel in the center of Vienna during the rush hour. The purpose of the experiments was to try to influence the frequency of both the cars entering the tunnel and the persons entering the supermarket by different subjects who were provided with auditory feedback at a later time.

PROCEDURE

For experiments 1 through 3 a photo-electric device was

installed at the entrance of a Vienna supermarket by a collaborator who did not know about the purpose of the device. The device produced a click in an intercom every time when interrupted by a person entering the supermarket, and a second click, when the person had already passed the light-beam. Persons did not know about the purpose of the light-beam and could not hear the clicks when passing through it, since the clicking intercom was placed in an adjoining room in an acoustically shielded box together with a microphone which recorded the click sequences.

For experiments 4 through 6 the photo-electric device was placed at the end of the tunnel. This tunnel can only be passed by one car at one time; likewise, only one person could enter the supermarket at one time passing the photo-electric device. No person is able to walk through the very small tunnel. The intercom for the photo-electric device placed at the tunnel was located in a room in the house above the tunnel, where I was allowed to use a small almost unused storage area in which to place the clicking intercom in the acoustically-shielded box, again together with the microphone. The installment and recording for all six experiments was performed by a person who did not know the purpose of the experiments. Moreover, he was not allowed to be present during the recording of the clicks in order to avoid any possible feedback. Taking into account technical problems for simultaneous feedback from these REG's the study was completely designed as a time-displacement study.

The recordings which were made during several weeks in the winter of 1978/79 were used for feedback one to two and a half months later.

Experimental set-up:

The experimenter monitoring the tape equipment was placed in one room and the subject given feedback in another room separated by a third room in between providing acoustical shielding. None of the subjects were aware that the click sequences played back to them via the headphones were pre-recorded. Instead they thought that they had to try to PK influence simultaneously generated clicks. Subjects were seated in comfortable reclining chairs and asked to listen to clicks played back through the headphones. They were asked to concentrate on the clicks in a way they thought

would be best in order to enhance the click-rate and to maintain the imagination of a great number of persons entering the supermarket or a great number of cars passing through the tunnel. In order to relax both experimenter and subject, a short speeded-up movie film was shown to the subject showing the persons in the supermarket and the cars coming out of the tunnel. Usually, the subjects found this amusing and wanted to try to visualize the quick-motion of persons or cars for their own PK efforts. Before the experiment started, the individual volume for the loudness of the clicks was pre-established together with the subject and a short exercise of progressive relaxation was provided while telling the subject that relaxation during psi experiments is very helpful for getting psi-hitting. Moreover, the subject was told that there would be ten periods of 40 seconds each, during which clicks could be heard through the headphones and ten 40 second intervals during which no feedback would be provided (control periods) and during which they should try not to think about the persons or the cars and simply relax. After each of the twenty intervals, a twenty-second intervening rest period followed at the end of which a one-second buzzer sound was played back to the subject if a conformance period was to follow and he or she should get ready to concentrate again on enhancing the click-rate, and without buzzer if there was a control period to follow the rest period. As during the control periods, no feedback was provided for the intervening rest period.

After the subject had put on the headset, the experimenter left the room and went into the monitoring room. There he could not hear the clicks and had no feedback whatsoever regarding the pre-recorded events. Nonetheless as an additional acoustic shielding, the experimenter was provided with another set of headphones through which he could hear pink noise during the whole run of the experimental sessions. He sat in front of the cassette-deck on which the clicks produced by the persons or the cars were recorded. The tape-recorder was shielded visually from the experimenter in order to avoid visual feedback by watching the VU-meters display. He was furthermore provided with a device with three buttons constructed for previous experiments (Gruber, 1979a; 1979b), one for the switching on of the tape, one for adjusting the volume of the subject's headphones, and one for the production of the one-second

buzzer sound also fed through the subject's headset and indicating the start of an experimental period. The experimenter was furthermore provided with a digital watch, which started working at the onset of the tape-recording and with a pack of twenty cards containing ten black and ten white cards. Immediately before each session, the experimenter shuffled the pack of twenty cards twenty times and placed the shuffled pack upside down in front of him. By taking a card from the pack before each 40-second interval, it was determined whether an experimental or a control period was to follow (a white card indicated an experimental period, a black card, a control period). It was decided to use the cards as a randomization procedure in order to have an equal amount of control and experimental periods. Each time an experimental period was to follow the rest period, the experimenter provided the one-second buzzer sound immediately before its commencement. He then also pressed the button for the switching on of the volume in the subject's headphones. When the volume was turned on, a control lamp was lit which could be observed by the experimenter and which was turned off with the turning off of the subject's headset volume. At the end of an experimental session, three consecutive buzzer sounds were provided which was the signal for the subject to take off the headset and to leave the experimental room without further contact with the experimenter. This was done, as in previous experiments, in order to avoid sensory cueing about possible beliefs of the subjects about their alleged success or failure, which could lead to some sort of secondary feedback on the experimenter and subsequent psi experimenter influence on the outcome of the experiments.

The name of the subject, as well as the sequence of intervals was noted for each tape. The co-experimenter who did not know about the experiments and who installed the photo-electric devices was asked to check the click-sequences according to the control and experimental periods and to perform a statistical analysis in a prescribed way. Results were re-checked by another person not aware of the experiments.

It should be mentioned that I first had in mind recording together with the click-sequences, the noise produced in the supermarket and that of the cars, in order to facilitate a vivid imagination of the "REG" for the subjects. I listened to such preliminary tapes and it was often hard to hear the clicks, since the environmental noises differed very much

in loudness. There might also have been problems in counting the clicks afterwards since there might have been noises on the tapes similar to the clicks from the photo-electric device which would have been erroneously interpreted as clicks experimentally produced.

RESULTS

It was decided to calculate a Mann-Whitney U between the 10 experimental and the 10 control periods for each subject or session respectively. The probability values thus derived were combined across subjects by the chi-square method recommended by Guilford (1956). One-tailed tests were chosen since similar experiments as well as a short pilot study yielded positive results in the predicted direction. The hypothesis was that there would be significantly more clicks during the experimental periods than during the control periods for each group of single sessions forming one experiment and for overall results. As significance level the 5% level was chosen.

Experiments 1 through 3

During experiments 1 through 3, subjects listened during the experimental periods to pre-recorded click sequences produced by persons entering a supermarket. In experiment 1 a special subject, who was especially motivated to participate in the study and who had shown good results in a similar, previous study (Gruber 1979a), contributed 10 sessions yielding a non significant p-value of .30 (cf. table 1). Experiment two was performed by ten unselected volunteer subjects each contributing one session, and for experiment three, again ten unselected subjects each contributed one session. Only experiment two yielded a significant p-value ($p = .05$). In spite of having only one significant result in the first three experiments, the results of experiment one and three had a tendency in the predicted direction thus yielding an overall positive and significant result for experiments 1 through 3 and corroborating with the second hypothesis (overall $p = .05$). Thus the behavior of a group of people in a real-life setting has been influenced by motivated distant subjects at a later time.

TABLE 1
Results of experiments 1 through 3 (supermarket)

Experiment	Chi-square	df	p-value
(1)	23.21	20	.30
(2)	33.94	20	.05
(3)	23.67	20	.30
total	80.82	60	.05

all p-values one-tailed

Experiments 4 through 6

During experiments 4 through 6, subjects listened to click-sequences produced by cars passing through a tunnel during the experimental periods. In experiment 4 the same special subject as in experiment 1 again contributed 10 sessions. The same subjects as for experiments two and three participated in experiments five and six, but in a different order than in the previous experiments. The special subject ($p = .01$) as well as the second set of subjects ($p = .05$) produced significant results in the predicted direction yielding a significant overall result of $p = .005$.

TABLE 2
Results of experiments 4 through 6 (tunnel)

Experiment	Chi-square	df	p-value
(4)	39.35	20	.01
(5)	20.49	20	.50
(6)	33.91	20	.05
total	93.75	60	.005

all p-values one-tailed

Thus the behavior of a group of people driving cars in a real-life situation has been influenced by motivated distant subjects at a later time.

Subjects did not score consistently lower or higher during the experimental periods than during control periods. Moreover, the results were not uniformly distributed across subjects. Statistical analysis of these features were not performed since they did not serve the purpose and hypotheses of the study.

DISCUSSION

The reported experiments were partly able to corroborate with the first hypothesis and totally with the second hypothesis. Subjects were able to show an overall time-displaced PK effect on the behavior of a group of people in real-life settings who were not aware of participating in an experiment at all. The positive results in this study as well as in previous similar studies (Braud 1979a; Gruber 1979a; 1979b) could be associated with the fact that nobody, neither the subjects nor the experimenter, knew how many clicks were to be expected by "chance", making the system more flexible and open for PK influence (this suggestion was made by Carl Sargent during the discussion of my paper presented to the 1979 SPR conference in Edinburgh, Gruber 1979a). Moreover, the subjects were very motivated to participate in this study since a lot of them expressed particular interest in the "strange" task they had to accomplish. I must admit that I myself was very motivated and especially enthusiastic about this study too, because of the good results of my previous similar experiments and of the challenge of this study. Hence, there still might be the possibility of a psi experimenter influence, although I tried to avoid this influence as far as possible, at least concerning experimenter feedback during the experiment and data analysis. Persons other than the subjects observing the data were the two persons involved in checking the data and performing the statistical analysis. Neither of them were aware of the "meaning" of the data since they had no information whatsoever about the experiment. This feature together with the subjects being aware of the meaning of their observations and the fact that the expectation of pure "chance" in these experiments was not known, may account for an interpretation in the framework of the system-theoretical approach to psychokinesis by Lucadou and Kornwachs (1977; 1979).

In this concept, the less PK there is to be expected the more a subject (or person involved in the experiment) knows about the whole situation. The more one knows about the system in question, the more one is likely to be inhibited towards paranormal action. Since, in the present experiment it is to a great extent not possible to describe the system in a deterministic way, it is therefore open to indeterministic behavior.

On the other hand, these experiments, seen in connection with the previous experiments mentioned, can be conceived within the framework of Stanford's conformance behavior model by fitting into the basic requirements for the occurrence of conformance behavior.

REFERENCES

Braud, W.G. 'Conformance behavior involving living systems', R.I.P. 1978, Scarecrow Press, Metuchen, 1979.

Gruber, E.R. 'A study of conformance behavior involving rats and mice', Paper presented at the 3rd SPR Conference, Edinburgh, 1979 (a).

Gruber, E.R. 'Conformance behavior involving animal and human subjects', E.J.P., 1979, 3, 1 (b).

Guilford, J. 'Fundamental statistics in psychology and education', MacGraw Hill, New York, 1956.

Lucadou, W.v. & Kornwachs, K. 'Fitting of PK results with a quantum mechanical model?', Paper presented at the Parascience conference, London, 1977.

Lucadou, W.v. & Kornwachs, K. 'Development of the system-theoretical approach to psychokinesis', Paper presented at the Parascience conference, London, 1979.

Stanford, R.G. 'Experimental psychokinesis: A review from diverse perspectives', In Wolman, B.B. (ed) Handbook of Parapsychology, Van Nostrand Reinhold, 1977, 324-381 (a).

Stanford, R.G. 'Conceptual frameworks of contemporary psi research', In Wolman, B.B. (ed) Handbook of Parapsychology, Van Nostrand Reinhold, New York, 1977, 823-838 (b).

Stanford, R.G. 'Toward reinterpreting psi events', J.A.S.P.R., 1978, 72, 197-214.

ACKNOWLEDGEMENTS

Thanks are due to all friends who made it possible for the recordings to be made and to the Parapsychology Foundation, Inc., New York, who partially supported this study by a generous grant.

Elmar R. Gruber
Institut fur Grenzgebiete der Psychologie und Psychohygiene
Eichhalde 12
7800 Freiburg
W-Germany

THE DISTRIBUTION OF PSI ABILITY AND PSI PERFORMANCE

Jürgen Keil

University of Tasmania

Pratt's reminder (1978) to question assumptions, general concern about experimenter effects (Kennedy & Taddonio, 1976; White, 1976), a recent followup study of 30 mini Geller cases (Keil, 1979) and a hitherto unpublished case of metal bending probably provide a suitable background against which Millar's radical views about psi sources (1978, p. 324) may be discussed.

The term psi ability may be defined as a relatively permanent disposition of a person to make under suitable conditions appropriate responses on the basis of which the involvement of psi can be inferred. For the purpose of this discussion it is of no importance whether this ability is regarded as being genetically determined or learned or is seen as being based on a combination of the two. However, it is assumed that this disposition is relatively permanent by the time a person can participate in a psi investigation. The strength of this disposition is assumed to differ from person to person and may follow something like a normal distribution. While it is difficult to measure psi ability, an indirect indication can be obtained in terms of certain kinds of maximum performance. A particular person may be able to move an object weighing x grams over a certain distance. It then seems reasonable to assume (although it may not necessarily be fully justified to do so) that in general this person will in principle over a long period of time (of the order of years), continue to be able to move x grams, again over a similar distance. Psi performance may be defined as the actual responses made by a person on the basis of which the involvement of psi can be inferred. Psi performance is normally evaluated over relatively short periods of time.

The two terms, ability and performance, are suggested because there seems to be evidence that in spite of a perhaps normally distributed ability persons do not perform in agreement with this distribution. Millar (1978, p. 324) suggested that the experimental literature identifies a minority of subjects and experimenters as psi sources and that in comparison others

provide little psi input into test situations. Significant performances of groups of unselected subjects are largely due to psi provided by those experimenters who are able to obtain significant results with such groups. If a selected subject provides the psi input, different experimenters are often able to obtain positive results. D.D. Home may be regarded as such an example (see Beloff, 1977) and more recently Stepanek (see Pratt, 1973).

If the complexity of psi is considered it seems inevitable that some experimenters may influence the test situation in more than psychological terms by providing some or all of the psi input required in an experiment, even if in terms of experimental design and traditional expectations the participating subjects are supposed to be the main psi sources. Millar did not indicate whether he referred to psi ability or performance. In one way he seemed to restrict himself to measures which suggest performance only but his reference to unsuccessful experimenters also implies that he expects psi ability to be distributed like a step function. The likelihood of this has some considerable bearing on practical strategies in experimentation and it seems of interest to outline some arguments and findings which may be interpreted as being in disagreement with Millar's assumptions and which give some support to the notion that psi ability may be distributed normally.

Directly observable PK was quite rare until the appearance of Uri Geller in the early seventies. The reported Mini Geller phenomena which were frequently recorded in Germany (Bender et al., 1975; 1976a, 1976b) suggest a wide distribution of psi ability and seem to support Batcheldor's views (1966; in press) that under suitable conditions unselected subjects can produce large scale PK results.

However two questions require discussion:
(1) What are the justifications for the acceptance of a percentage of the mini Geller cases as indicators of psi?
(2) Could the percentage of psi mini Geller cases (that is, those cases in which psi is assumed to have been involved in the reported phenomena) be regarded as a similar minority to the one which according to Millar is responsible for successful psi experiments? Members of such minorities ought to obtain psi results relatively frequently.

In reply to the first question but also with some relevance to the second it may be appropriate to outline the P.F. case

which refers to events in October 1977. At that time Geller visited Hobart (population approximately 130.000) Tasmania, Australia for the first time and gave a public performance. In an interview arranged by the local television station, Geller spoke for approximately 4 minutes and included a brief suggestion that viewers should try to obtain paranormal phenomena such as metal bending. In contrast to some other similar media exposures, Geller did not spend much time (actually only 15 seconds) in connection with his suggestion for paranormal events. This author was able to obtain a copy of the videotaped Geller interview and also checked with the television station whether anyone had reported that metal bending took place. The station had only received two replies. One referred to a boy whose parents did not wish any further discussion or investigation of the events. The second reply came from P.F. This author had also heard of additional cases in the community who were however even more concerned not to be involved in any further discussions or investigations.

P.F. is a professional man, married with seven children. He had not been interested in psi before and when he saw Geller on his television screen in his home he would not (as he related to this author) have stood up to fetch a spoon. However, since one was within reach on a table, P.F. picked it up and held it in his right hand and to his amazement and surprise the spoon clearly bent by about 70 degrees. The bending took place about 50 to 60 mm from the nearest point at which the spoon was held. P.F. simply held the spoon at its end, dit not touch it in any way with his other hand and did not carry out any other manipulations. The bending process seemed to be a continuous one and lasted for about three to four seconds. The end product was a clearly bent spoon which could not have mistakenly been regarded as a straight spoon which was thought to have bent on that occasion but which was really bent from the start. During these events P.F.'s mother-in-law was in the same room but did not observe the spoon as she was sitting closer to the television set than P.F. She was equally amazed when confronted with the spoon and it seems reasonable to rule out the possibility that P.F. made a mistake of the kind as suggested above.

P.F. had no hesitation to talk to this author. At the same time P.F. was in no way interested in publicity and except for a small circle of friends and relatives there is no public awareness of these events. P.F. rang the station expecting that many others had rung as well and he simply wanted to add his experience to what he expected to be a number of

similar occurrences. When P.F. realised that this author had a professional interest in such events he agreed to participate in further attempts which might lead to repeated demonstrations and hopefully to experimentation. P.F. had no difficulty in accepting the view that the metal bending was primarily due to his own psi rather than being directly due to Geller.

P.F. tried repeatedly to obtain similar effects while observing Geller and listening to Geller on a video recording of the same interview. These attempts were carried out by P.F. on his own in a university room where the video equipment could be set up. When P.F. remained unsuccessful under these conditions this author made a super 8 cinefilm copy of the video tape and provided P.F. with this copy and a sound projector to take home in order to attempt metal bending under similar environmental conditions. P.F. tried a number of times under circumstances (such as time of day) resembling as closely as possible the original circumstances but he had no success. Since the first occasion in October 1977 P.F. had not experienced anything suggestive of psi.

This case is of interest for two reasons:
(1) The probability that the 1977 experience of metal bending was real and indicative of psi, is very high indeed. This author as well as other members of the psychology department, University of Tasmania, had known P.F. for some years. It does not seem to be reasonable to assume that P.F. might have intentionally or unintentionally produced the effects by normal means. This is necessarily to some extent a subjective assessment which is nevertheless shared by my colleagues who know P.F. and who have no particular interest in psi phenomena. Many similar cases occurred in Europe earlier (Bender, et. al., 1975, 1976a, 1976b; Keil & Hill, 1975) and there seems to be a good justification to accept a substantial percentage as indicative of psi. In a recent investigation of 30 mini Geller cases (Keil, 1979) five of these were estimated to be indicative of psi phenomena.
(2) If it is accepted that Geller triggered a substantial number of psi phenomena it could still be argued that these phenomena were associated with the same kind of minority of subjects and experimenters who are responsible for significant psi results in experimental situations. The relatively large number could still be explained if it is realised that on account of the mass media more people than ever before were encouraged to experience psi. It would be reasonable to accept the proposition that most of the persons who reported psi phenomena and who also stated that they had not experienced anything paranormal before, simply had not tried to be involved in such an experience

previously. However, if they represent the able minority it would also be reasonable to expect that they would continue to have such experiences once they had subjectively accepted them as real.

The P.F. case showed that even under relatively similar circumstances this does not readily happen and the investigation of the 30 mini Geller cases indicated that although after four years the events were still regarded as real and paranormal, no further comparable paranormal events had been experienced since. Field experiments carried out by this author suggested that psi might still be detected but no directly observable events could be clearly based on psi.

There have been cases where apparent psi phenomena, initially triggered by Geller, continued for some time but it seems that a far greater number of people had only one experience which did not recur over a period of four years.

These observations rather suggest that the successful mini Geller cases represent an extension to Millar's minority of good subjects and experimenters. That is, even if the distribution of psi performance is in closer agreement with a step function than with a normal curve, psi ability may still be normally distributed. The main factors which determine whether a psi component can enter a performance task seem to be psychological ones. To specify such factors is likely to remain a difficult task for some considerable period of time. Nevertheless, Batcheldor's suggestion (1966; in press) could provide some hope of constructing a framework within which the psi mini Geller cases may start to make some sense. According to Batcheldor two main conditions for psi performance are important and perhaps essential:
(1) A heightened involuntary belief state. A conscious desire for psi phenomena is not equivalent to this involuntary state. It is possible (with a substantial degree of success) to induce such states experimentally by creating environmental settings in which normal events may be experienced as paranormal ones.
(2) Irrespective of a conscious desire for psi phenomena many people have fears of such phenomena below the level of awareness. These fears have to be removed or reduced through a process which is not dissimilar to desensitisation.

It seems desirable to encourage Batcheldor's procedures on a larger scale and it is possible that some indications will emerge that the distribution of psi ability and perhaps even psi performance is following a normal curve after all.

NOTE

The psi distribution problem has recently been discussed in relation to statistical tests. No publications relating to these discussions were available when this paper was submitted and the arguments presented in this paper are primarily put forward in order to encourage research involving directly observable PK.

REFERENCES

Batcheldor, K.J. 'Report on a case of table levitation and associated phenomena', Journal A.S.P.R., 1966, 43, 339-356.

Batcheldor, K.J. 'PK in sitter groups', Psychoenergetic Systems, in press.

Beloff, J. 'Historical overview', In Wolman B.B. (ed.), Handbook of parapsychology, New York: Van Nostrand Reinhold, 1977.

Bender, H., Hampel, R., Kury, H. & Wendlandt, S. 'Der "Geller-Effekt" I', Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie, 1975, 17, 4, 219-240.

Bender, H., Hampel, R., Kury, H. & Wendlandt, S. 'Der "Geller-Effekt" II', Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie, 1976, 18, 1, 1-20.

Bender, H., Hampel, R., Kury, H. & Wendlandt, S. 'Der "Geller-Effekt" III', Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie, 1976, 18, 2, 105-116.

Keil, H.H.J. 'Field experiments with 30 possible PK subjects', E.J.P., 1979, 3, 1, 21-36.

Keil, H.H.J. & Hill, S. 'Mini Geller PK cases', In Morris et. al., (eds) Research in Parapsychology 1974, 1975, 69-70.

Kennedy, J.E. & Taddonio, J.L. 'Experimenter effects in parapsychological research', J.P., 1976, 40, 1-33.

Millar, B. 'The observational theories: a primer', E.J.P., 1978, 2, 3, 304-332.

Pratt, J.G. 'A decade of research with a selected ESP subject: An overview and reappraisal of the work with Pavel Stepanek', Proc. A.S.P.R., 30, 1973, 1-78.

Pratt, J.G. 'Prologue to a debate: Some assumptions relevant to research in parapsychology', J.A.S.P.R., 1978, 72, 127-131.

White, R.A. 'The limits of experimenter influence on psi test results: Can any be set?', J.A.S.P.R., 1976, 70, 333-369.

H.H.J. Keil
University of Tasmania
G.P.O. Box 252 C
Hobart, Tasmania
Australia 7001

COMMUNICATION AND IMAGERY RESONANCE: A THEORETICAL APPROACH TO PSYCHIC HEALING

Ralph G. Locke

Visiting Professor, Soc. Dept.
University of North Carolina
Chapell Hill

Resonance is most commonly understood as a purely physical phenomenon. However, in the last decade it has become associated with some profound thinking in human communication studies. This line of thought, beginning with semiology, shows some promise as a framework for investigating psychic healing. In particular, resonance gains substance when it is related to language and imagery processes. Recent innovations in psychotherapy and allopathic medicine which utilize imagery for the purposes of symptom management and more profound physical and psychological adjustment allow an expansion of the imagery notion. Specifically, imagery is considered to be an information storage and adaptation system of the human organism. Under certain conditions of senders (healers) and receivers (patients) of communication, these systems may be changed to effect therapeutic movement. The theoretical approach offered here extends these concerns into four major areas of contemporary research and relates them to the problems and possibilities in the explanation of healing and psi. There are four approaches to resonance.

1. SEMIOLOGY

In general, semiology is concerned with the study of human communication, but in a particular way. Its focus is upon relations of signification, expressed in the correlation S/s. Here, S and s are two

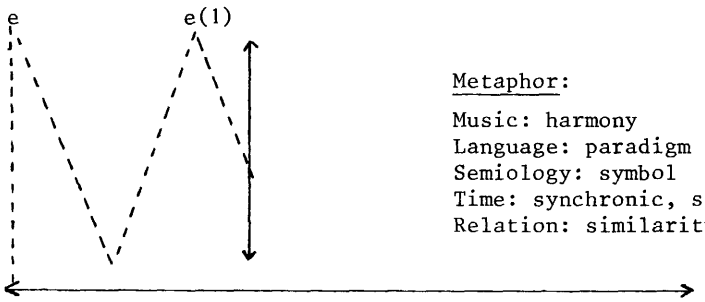
complementary entities: S is the significans and s is the significatum. So, the sign relationship serves to correlate a formal content with its expression. Moreover, the kind and range of the correlations covers the spectrum of human communication and consciousness. Semiology, then, is concerned with symbolic function and asserts that this function defines human thought and expression (Leach, 1976). In the terms set out, semiology is usually identified with structuralism (Broekman, 1974).

Superficially, semiology seems to offer no new insights into healing processes: there are several other well-developed communication theories of healing. However, the key to semiology is in its notion of levels of structure. In this, there is a dialectical relationship between the surface structure (manifest pattern and content) of communication and consciousness and a deep structure (latent pattern and configuration). The most striking example of this lies in the relationship between metaphor and metonymy. Within communication we can identify recurrent patterns of association which vary in the extent to which they are conventionalized. This is the axis of metaphor which is synchronic and based on the relation (correlation) of similarity. On the other hand, elements of expression are understood in another way: 'sense' is achieved out of the relation elements have to contexts, or perceptual and imagery configurations. The relationship of metaphor to metonymy can be grasped most clearly in the example of an orchestral performance. Listening in one way, the association of separate instruments can be discriminated as harmony, while a change of focus shows us the temporal unfolding (diachronic) of melody (configuration, context). But, routinely, we perceive the performance in a unitary sense just as we understand dialogue. Figure 1 sets out some aspects of the metaphor-metonymy relation (see Wilden 1972: 55).

In short, human communication, whether linguistic, or bodily expression (e(1), e(2), . . ., e(n); see figure 1) is made sense of by its successive contextings, so that the relationship metaphor/metonymy is equivalent to digital/analogue and verbal-logical/imagery configuration (Wilden, 1972; Singer, 1978). Resonance analysis in semiology follows in this general frame of reference. Three brief discussions of semiological applications of resonance follow.

A. Jacques Lacan:

This first example deals with the application of semiology to psychoanalysis as found in Lacan's (1968; 1977a; 1977b) work. Lacan



Metaphor:

Music: harmony
Language: paradigm
Semiology: symbol
Time: synchronic, simultaneity
Relation: similarity, association

Metonymy:

Music: melody
Language: syntagm
Semiology: sign
Time: diachronic, succession
Relation: part-whole, element-configuration

e = element
(of discourse
harmony, action)

Figure 1

offers a reformulation of psychoanalysis, suggesting that the unconscious has something of the properties of a language. In outline, his argument concerning this and the method of 'analysis' he deems appropriate to it have the following components.

(i) The connection between S and s is a function of an arbitrary linguistic code. Signifiers are connected in a linguistic space, forming a signifying chain. Each chain has two dimensions - - a set of cultural conventions transcending subjectivity, and a set of unique, personal significations. Therefore, context has these latter two dimensions.
(ii) The connection between signifiers is linear and vertical - - i.e., metonymic and metaphoric, as set out in figure 2.
(iii) In speech, particular words are nodes for a particularly salient and polyphonic chain of signifiers. The therapist's task is to achieve a state of resonance with nodal points in the patient's discourse.
(iv) Resonance depends upon letting oneself be dominated and controlled by a word to which one must correspond and listen (Gorney, 1978: 255) as opposed to controlling and speaking directively.

(v) This receptive attitude, a passive 'listening', allows the therapist to then speak evocatively - - i.e., to elicit the hidden pattern of the patient's speech. However, this pattern is in some sense 'known' by the patient and it contains his authentic or genuine self.

> That is to say, what the broadcaster, the therapist, is saying is not something new to the patient, but resonates with something the patient already knows and brings it into clearer focus. The concept of resonance permits one to explore the possibility that change results as a consequence of the expansion of configurational patterns over time (Levenson, 1976: 13).

(vi) The next step in the sequence, to achieve the 'expansion' noted, is to use the code of patient's inner language to play back their messages to the therapist in inverted form. Configurational change amounts to 'insight' in the usual psychoanalytic setting.

Lacan's approach to resonance rests very squarely upon the adoption by the therapist of an appropriate attitude. It is not completely described by 'empathy'; rather, Lacan's description is more like a phenomenological attitude where one attempts to 'see through' (or 'hear through') the 'cloud of ideas' (Marcuse, 1974: 232). It is a 'letting go' of one's habitual style of perceiving - - an altaration of taken-for-granted boundaries implicit in perception. James Gorney (1978: 26) gives an example of this kind of listening, relating to his treatment of a schizophrenic patient.

> As I let her words flow over me, I began to notice periodic experiences of my head seeming to swell, of my mind seeming to be on the verge of leaving my body, and occasionally there were thoughts that I might be about to go crazy myself and start screaming and babbling. These experiences were quite alien to me.

Gorney's experience was one of assimilating the code of the patient's communication, so that the configurations in it initially produced disorientation. Boundary alteration of this kind with attendant 'symptoms' is found not only in schizophrenia (Bateson, 1978; Berger, 1978), but also in mediumship (Locke, 1979), healing (Douglass, 1969), and culture shock (Castaneda, 1969). Moreover, it has a strong similarity to the 'tuning in' process common to several humanistic therapies and communion involved in some psychic healing methods (LeShan, 1974, 1975). In short, it is a kind of phenomenological attitude (Van Kaam, 1959; Zinker, 1977).

In the case reported by Gorney, and also those referred to by Bateson and Berger, boundary changes occur when there is a 'splitting' of the referential (signification) structure of communication. Resonance cannot be achieved by hearing only one strand of meaning.

A complementary viewpoint on healing is provided by another French scholar, Claude Levi-Strauss, whose approach is also structural-semiological.

B. Claude Levi-Strauss:

A single representative and extremely important example of Levi-Strauss' analysis will be dealt with here - - shamanistic cures amongst the Cuna Indians of Panama. Under the heading of 'the effectiveness of symbols', Levi-Strauss (1963) discusses the shamanistic management of difficult childbirth through the use of a traditional, mythically organized song. At the outset, Levi-Strauss makes it clear that his interpretation of the song is that it constitutes a psychological treatment, but points out that this does not make sense unless we can show how 'specific psychological representations' (ibid, 191) impinge upon definite psychological difficulties. The song is involved only in cases of difficult childbirth, where the woman experiences delay and pain. It involves the following features:

(i) The shaman speaks to the patient, interpreting her labor pains and anatomy in terms of the mythical composition of the song. A transformation of boundary relations takes place: the external, physical world becomes physiological and the mythical world becomes psychological. Mechanical-anatomical aspects of labor are translated into metaphors of the myth.
(ii) When the shaman speaks, he speaks to something already known by the patient - - the myth and the cultural conventions which surround it. Pain not previously accepted, a confusing and obscure situation, is made meaningful by the myth. In other words, the myth provides a language whereby these previously ineffable states can be expressed.

> And it is the transition to this verbal expression - - at the same time making it possible to undergo in an ordered and intelligible form a real experience that would otherwise be chaotic and inexpressible - - which induces the release of the physiological process, that is, the reorganization, in a favourable direction, of the process to which the sick woman is subjected (ibid, 198).

(iii) The shaman leads the patient through a mythical landscape, matching a physiological topography, to successful childbirth. By so doing, he is 'mythically transmuted' and the patient identifies with him. Successful healing, then, depends upon accumulations of experience of certain kinds so that a healer can address himself to them and produce an abreaction - - an emotional reliving - - through his resonance. For the Cuna, the abreaction is in terms of the emotional qualities of the myth; for members of our society, abreaction falls into the domain 'personal myths' and stories which surround the pathological nucleus (May, 1973).
(iv) Levi-Strauss takes some pains to include the full-range of communication in his study and theoretical position. In particular, effective healing is defined in terms of a range of ritual and habitual actions, including speech forms, which symbolize elements of the situationally described problem.
(v) A final but very significant proposition made by Levi-Strauss is that we are quite possibly dealing with homologous structures in the process described. At varying levels - - biochemical, physiological, conscious, unconscious - - we may find structural identities which are the keys to the transformability of one epistemic state into another - - e.g., mythical (social) --> psychological --> physiological. Two additional points are relevant here: (a) Content may vary within and between cultures, especially with respect to the three states just mentioned, but it is the structure of consciousness framing content differences which is important. So, resonance may not merely be a matter of content-matching but may necessarily be a relation of homology. (b) The Cuna process of effective healing seems to involve an inversion of certain elements defining the problem, as in Lacan's strategy. Broadly stating it, the inversion is of the form:

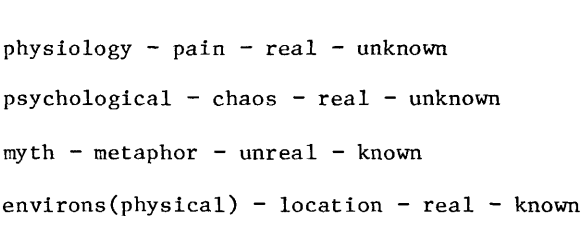


Figure 2

Transformations are achieved by the shaman's creation of resonant contexts (metonymy) in which metaphors and the corresponding structures contexting them can be understood. Just as in Lacan's approach, metaphors are the parts which evokes the wholes - - configurations which are imagistic and epistemic in the sense of providing continuous coding frames for discrete events.

Levi-Strauss' position is distinctive for its extension of the healing relationship beyond psychological considerations to suggest a complex formula for transformation of social and communicative processes into physiological effects. Recent imagery techniques in psychotherapy deal with these transformations in a particular way. Specifically, elicitation of fantasy-imagery sequences, of a free associational kind or guided according to some formula, serves a therapeutic model. For instance, Guided Affective Imagery (GAI) has treatment goals stated as resymbolization of experience and symptom relief (Leuner, 1969). Looking at the three principal neural encoding systems as set out by Horowitz (1977), verbal-sequential (lexical), motor-kinetic (enactive), and visual-auditory-spatial (image), we derive the following picture of GAI.

System 1 (lexical) --> System 1 and/or System 3 (image) --> System 2 (enactive).

In several healing techniques there is an emphasis upon similar patterns where healer-patient interaction builds up a mutually held knowledge of the patient's situationally evoked imagery experience. To some extent, this knowledge is equivalent in function to the mythical domain in Levi-Strauss' example, insofar as it becomes the instrument of resonance. Parallels exist in group therapy and in esoteric healing systems which use regression analysis (ego-state regression and past lives recall), and divination of hidden orders of patients' lives which are reified and become, again, the instrument of resonance (Lauer, 1973; Macklin, 1972; Prince, 1973).

C. Bandler and Grinder:

This last of the semiological approaches is important because of its emphasis upon the use of language in mapping patients' epistemic

states, which Bandler and Grinder (1976) refer to as 'representational systems', and in producing change. Strictly speaking, Bandler and Grinder's material does not fit into semiology. However, it draws upon structural linguistics and uses signification and imagery processes as the cornerstone of treatment. Aspects directly involved with resonance are the following.

(i) Human beings do not apprehend 'reality' directly. They construct models or representations of it. Characteristically, communication and action is in terms of representations, so that individuals expressing pain and dissatisfaction (e.g., presenting themselves for psycho-therapy) typically have 'problems' with their representation systems.
(ii) Representations are shaped, as in mapping (a meta-model), by the major input channels responsible for information reception - - visual, auditory, kinaesthetic. Other sensory modalities are regarded as less important in representation-building.
(iii) Information received in one channel may be stored in another. Visual information, for example, may be stored in a lexical code and communicated in natural language terms. However, despite this mode-switching, it is possible to detect the representation system supplying the background configuration to communication by reference to consistent use of representation-specific terms (metaphors in semiology). In other words, digital representation is used to indicate and divulge analogue representation.
(iv) Representation systems are differentially valued so that individuals' habitual communication style reveals their most valued representations (MVR). Establishing rapport - - empathic and apparently genuine communication - - depends upon the healer's (usually a psychotherapist in examples cited) correctly assessing the patient's MVR and using it in his own language. That is, the healer initiates a receptive healing set in the patient by speaking in terms already 'known' to the patient; in other words, imagery matching or resonance. Where a patient consistently uses kinaesthetic predicates (e.g., feel, sense, touch), the healer uses kinaesthetic predicates.
(v) Bandler and Grinder extend their analysis by posing this basic set of ideas along with Bateson's (1972) theory of logical types in communication and, more importantly, including this theory of modal confusion and contradiction as the genesis of adjustment problems. This greatly refines the treatment of how modal dominance (MVR) is detected, and shows something of how difficult it is.

Bandler and Grinder's attempt to reconstruct psychotherapy is important for its success in intergrating transformational grammar,

communication ecology, and imagery analysis and providing concrete procedures for effective treatment. In general terms, Bandler and Grinder offer a theoretical analysis which is continuous with those of Lacan and Levi-Strauss in underlining elements essential for resonance to occur - - viz., a particular attitude of the healer, a certain style of communication, and the centrality of part-whole relations (metaphor/metonymy) in the production of insight and (physical and psychological) change. The latter emphasis has been carried forward recently by Gordon (1978).

2. INTERPERSONAL PSYCHOLOGY

Undoubtedly, the first section on semiology is interpersonal, but this part differs in that its focus is more upon the establishment and management of the social-interactional styles central to psychoanalysis and some humanistic psychotherapies.

A very large area of research in psychotherapy is that concerned with looking at the essential ingredients of effective therapy considered as one kind of helping relationship (Garfield and Bergin, 1978; Gurman and Razin, 1977). Much of the research done revolves about the issue of what attitudinal requirements are necessary for the therapist to establish a positive relationship with the patient, and, in addition, what is necessary to carry this relationship forward to a more constructive adaptation for the patient. Pioneering work in this area was done by Rogers (1957), and Rogers and Truax (1967). Rogers' initial position was to argue for "the sufficiency of the therapeutic conditions of successful initiation of a relationship, but are insufficient, separately or together, to guarantee positive outcomes. However, it seems clear that these variables are relatively independent aspects of the quality of the therapist's 'basic feelings' or 'underlying attitudes'" (Rogers, 1967). Certainly, a continuing difficulty in assessing the attitude-outcome relationship is dealing with the troublesome notion of empathy.

Rogers (1957) and Rogers and Truax (1967) deal with this part of the therapeutic relationship by distinguishing between empathy as the general process of 'feeling with' and understanding (verstehen) and accurate empathy which is described theoretically as the 'ability of the therapist accurately and sensitively to understand experiences and feelings and their meaning to the client during the moment-to-moment encounter of psychotherapy'. Thus, accurate empathy is not technical, diagnostic mastery, but an expression of a phenomenological attitude. It is a form of 'seeing', 'listening', and 'feeling' from the

standpoint of the other, so that, as far as is possible, one's habitual communication and perceptual sets do not filter out what the other is saying and doing. According to Gurman (op.cit.: 504), "empathy (accurate empathy) goes beyond empathic inference - - that is, a cognitive-intellectual understanding of the patient - - to empathic communication 'in a language attuned to the client's current feelings'".

A more extensive discussion and analysis of resonance is undertaken by Watkins (1978). His position is basically a variety of psychoanalysis. In approaching the concept of resonance, Watkins begins by giving the example of physical (acoustic) resonance in musical instruments, and then moves to psychological and interpersonal parallels. Here, he distinguishes between resonance and identification. Watkins sees resonance as a temporary process for the purposes of inducing therapeutic movement, while identification is associated with more permanent changes. However, resonance is a temporary form of identification. Moreover, Watkins defines it as a psychological resonance, and, although he develops a method of scoring therapists' resonance ability in interaction with patients, his treatment is predominantly psychodynamic. For Watkins, resonance is, "that inner experience within the therapist during which he co-feels (co-enjoys, co-suffers) and co-understands with his patient, though in mini-form" (ibid: 234). Why Watkins uses 'mini-form' is not made clear. However, we can speculate that it follows from his distinction between resonance and identification. Namely, identification is a complete matching of experiential structures while resonance is not, since it is necessary for the therapist to accurately empathize, self-consciously monitor interaction, and produce resonant actions (Orlinsky and Howard, 1978: 317).

In summary, then, the positions represented so far set out certain features of resonance. Firstly, resonance is not simply a matter of accurate empathy, but also of empathic communication. Secondly, resonance falls between identification and alienation, so that it involves neither the complete 'immersion' in the other, or 'surrender' to the other, nor treatment of the other as something object-like and the product of impersonal, diagnosable forces. In the third place, resonance is achieved through adaption of a particular attitude of the healer which qualifies the ego-centric (and culturally transmitted) boundaries of 'seeing', 'feeling', and 'hearing'. The essence of this style is to suppress, or hold in abeyance, the logical and linear (lexical/digital) and to allow the configurational (imagistic/analogue) to predominate in relations of healing.

3. NEUROPSYCHOLOGICAL

In two recent, provocative articles, Edgar Levenson (1976, 1978) attempts to unravel the enigma of how therapeutic change occurs and offers an innovative way of theorizing about it. The framework of the argument is the familiar one: Is psychoanalysis an art or a science? Levenson discusses this issue and surprisingly comes up with a synthetic viewpoint derived from his survey of several fields, and, in particular, applications of hologram physics and mathematics.

> -- the model of neuropsychological functioning, the model of the therapeutic engagement, the model of intervention and the model of insight and change are all truly homologous; that is, they are connected not by analogy or metaphor but by real similarity and isomorphism in structure (1976: 9).

After a fashion, Levenson supports Levi-Strauss' theoretical position on levels of relationship and connections between structures. What leads Levenson to this position is an examination of the healer-patient relationship in terms of the knowledge which is gained and used in therapy. Specifically, using psychoanalytic theory, he suggests that each moment of the patient's discourse contains in it the configuration which contexts it and gives it meaning. In his 1978 work, Levenson interprets this part-whole relation in terms of metaphor and metonymy, digital and analogue processes, speculating a connection with left- and right-hemisphere functions, respectively.

What the analyst does is to learn the connections between moments of patients' actions and the configurations (epistemic systems) underlying them, so that, once he has a grasp of this, and it is done out of an open mutuality with the patient, he can then use it to produce resonance. In this he draws upon Lacan, asserting that at the conclusion of therapy the patient will have learned what the therapist knows, but this will be nothing that he already did not know. How the patient communicates what he knows and the content of what he knows is learned by the therapist and played back to him. Again, this must take place in a setting of trust, and it requires a certain attitude, Levenson (1978: 62) concludes, "There is a Sufi saying: 'Don't learn, listen.' It's not bad advice, but it's not easy either".

The curiosity of healing, though, is how the part, perhaps some minute physical adjustment as in Rolfian treatment, for example, produces radical changes in insight and, perhaps, integration of that insight into behavior changes. Levenson offers an explanation using

the model of a hologram. Very briefly, a hologram is produced when a single, laser light-beam is split so that one half illuminates a photographic plate directly, while the other half is reflected off the object to be 'photographed', resulting in a light interference pattern being set up on the plate. When a laser beam is projected through the plate, or any part of it, a three-dimensionsal image is produced in its entirity. As the piece of plate illuminated gets smaller, only details of the image are blurred. Levenson (1976: 16) notes, by way of indicating the paradigmatic significance of this, "So that the lens was the model of partial objects, the hologram is the model of whole objects, total objects that are invisible".

Cautious optimism suggests that the hologram may be the pattern of neuropsychological functioning so that some of the paradoxes of location (e.g., in memory studies) as opposed to diffuse or non-localized, atemporal, aspatial characteristics of the hologram are suggestive of connections with psi and memory (Roll, 1966) and a recasting of the synchronicity hypothesis (Jung, 1955; Pribram, 1971, 1979). Levenson goes on to propose that the logic of hologram operation may offer a fruitful view of resonance and cites Karl Pribram (ibid: 370) to provide a neuropsychological perspective.

> According to this view, thought is a search through the distributed holographic memory for resolution of uncertainty, i.e., for acquisition of relevant information. This formualtion is inadequate, however, unless the term relevant information includes appropriate configurations as well as items or bits in the information-theoretic sense. More often than not, when problems generate thought, contextual and configurational matchings are sought, not just specific items of information. These matchings, I believe, can occur best while the coding operation is in its holographic mode. Perhaps the power of thought in problem solving resides in the repeated return to the configurational form of representation that serves a rehearsal function and allows the occurrence of additional distributions in memory.

A related, though much earlier, point of view on resonance and neuropsychology is presented by Marshall (1960) who links resonance with psi. Marshall claims that his approach is 'holistic' rather than atomistic (digital) and it is cast in terms of structural similarity, structure-producing effects ('eidopoic influences'), and indeterminacy in complex (neuropsychological, in particular) structures.

Marshall draws his theoretical framework from several sources, but

principally from quantum theory and neurophysiology. The relevant application of his resonance theory here is in the area concerning indeterminacy in complex structures and eidopoic influence. In an analysis which has similar properties to that found in Markovian mathematics, Marshall sets out a relation between eidopoic influence, transitional probabilities, and structural change. Using the model of neuronal excitation (single-cells at first, then nerve-nets) he suggests that structural change only occurs when there is an indeterminate probability (between 0.1 and 0.9) in stimulus thresholds. In information-search and perceptual terms, a transition would be from one configuration (contexting digital stimuli) to another configuration as Pribram suggests. There is a 'critical area' in probability terms which allows the action of eidopoic influences.

There are several interesting features of Marshall's theory which have not been mentioned yet. Firstly, he notes that, 'Resonances appear as attractions between structures in pattern space; like vectors, they have direction and magnitude (ibid: 276). This is very much like the successive matching attempts in part-whole relations that are referred to by Levenson, Lacan, Levi-Straus, and Watkins. Secondly, Marshall puts forward a position strongly reminiscent of Merleau-Ponty's (1963) when the French philosopher attacked mechanistic psychologies. He says that his theory of resonance is more flexible than reflex or homeostatic mechanisms and it is more intelligent (as a model of human consciousness and behavior) "for it would respond more to generalized patterns than to exact details of the stimulus" (ibid: 278). This is, again, a call for a focus on part-whole relations and a concern for forms or structures rather than for linear-mechanical explanations. However, there are notable problems in these conceptions.

Where resonance is used as a theory of telepathy, then it has to account for eidopoic influences which are unconscious, and, correspondingly, for how information telepathically transmitted through imagery resonance can be recovered from the complex mass of imagery and perceptual data. Furthermore, there is some difficulty in discriminating which eidopoic influence is producing resonance - - i.e., resonance with several structures at once in the environment poses a profound problem of decipherment again.

4. PHENOMENOLOGY

It would be a gross injustice to attempt a treatment of 'phenomenology' in relation to resonance and healing; rather, it is

more faithful to this area of philosophy to say that we shall deal with 'phenomenologies' - - two out of the spectrum available.

a. Michael Polanyi

Undoubtedly, the focal concern of Polanyi is to understand the ways in which we 'know' the world, how we perceive it and act in it. His thesis is provocative insofar as it stands analytic philosophy and empiricism on their heads. He does this by posing two fundamental modes of knowing - - the physiognostic and the telegnostic (1969: 123-134) - - which stand in a from-to relationship. This is basically what makes his philosophy phenomenological: forms or wholes precede particular grasping (in language, perception, and tool use, for example). The paradigm for these assertions takes the form of the familiar process of pattern recognition common to Gestalt analysis - - e.g., recognition of a familiar face, a disease entity, or some other discernible entity. Polanyi points out that the physiognomy is apprehended, the whole recognized, so that the particulars of the whole are fused with it. In other words, we may denote specific features in the pattern, but the pattern is implicit in denotation - - i.e., the "joint meaning of all the instances to which the general (denotative) term is intended to refer" (ibid: 129).

Suitably, Polanyi refers to predominating physiognosis as 'tacit knowledge' (1964: 140). It is that which underlies our particular actions and yet is moulded to them in a relation strongly paralleling that suggested by Ryle (1976: 26-60) - - viz., 'knowing that', and 'knowing how'. The parallel is inexact, but the theme is continuous: a very large proportion of our actions involve implicit epistemic structures so that the 'knowing is in the doing'. Continuing on, while we can point to, reflect and theorize, these very 'actions' themselves assume a background of relevance. In general terms, Polanyi's distinction between kinds of knowing reflects the analogue-digital perceptual and communication organization which has been related to cerebral hemisphere differences (Bogen, 1973).

While this is an overly brief discussion of Polanyi's ideas, the purpose is to draw attention to levels of knowing in consciousness and action, and, to indicate that the physiognosis and telegnosis are dialectically related in our making 'sense' of anything. Finally, Polanyi (1964: 139) suggests that in the extent to which from-to relations are integrated may lie the key to creativity and directionality in consciousness. For he goes on to say that tacit knowing is vectorial or directional in the sense that it is

self-organizing, and in its directionality we find the potential for resonance as Marshall, Lacan, and Levi-Strauss point out.

b. Maurice Merleau-Ponty:

Stanley Krippner (1978: 72-73) strongly urges a wholism in dealing with healing, and underscores this when talking about psychic healing, "The physical and the psychological are two sides of the same coin; together they comprise the 'psychophysiological system'. A 'system' is an indivisible whole, one that loses its coherence when taken apart". From a rather different tradition, Merleau-Ponty (1962, 1963) emphasizes this thought. The French philosopher sets out to analyze the structure of perceptual experience, which task leads him to a rejection of Cartesian dualism and formulation of an alternative view. His original contribution arises out of a criticism of mechanistic psychologies and the notion that the organization of embodied consciousness is in terms of 'forms' or 'structures'.

> The form is a visible or sonorous configuration (or even a configuration which is prior to the distinction of the senses) in which the sensory value of each element is determined by its function in the whole and varies with it (1963: 168).

Part-whole themes are repeated here in conscious recognition and in unconscious levels, including the physiological and psychophysiological. Moreover, Merleau-Ponty is a 'moderate realist' just as is Polanyi in that he sees us in dialectical commerce with the environment, in what he calls an 'intentional arc'. What this latter implies is that our constructive engagement with the environment continues only insofar as the 'arc' or 'dialectic' is not broken. It can be broken between bodily and communicative surfaces and the environs as in schizophrenia and it can break down within the body surface as in physical trauma. Merleau-Ponty devotes many painstaking case-studies to this process of 'ruptured dialectics'.

What happens in some brain lesions, for example, is an interruption between epistemic structures. A striking example of this occurs in disorders where routine actions are 'forgotten' and the patient has to exert extreme effort to connect perception, intention, and action even for simple acts like picking up a pen. Each act has to be mentally rehearsed and self-consciously related to environs and the body and thus the 'intentional arc' is closed. In the introduction to the 'Structure of Behavior', John Wild notes of cases like this,

> Indeed, we find in the disintegrated consciousness an illustration of psychophysical parallelism, where conscious states run parallel to isolated bodily occurrences. In such sickness, this isolated body may causally affect my perception, so that what I perceive may lie like a 'subjective' veil between me and the real things and persons around me in the world (1963: xvi).

Starting from the position of 'the body as experienced', Merleau-Ponty demonstrates that in the sense of 'knowing how', our body 'inhabits' space, filling it with variously realized potentials for action. Embodied consciousness, the body image in particular, is an epistemic state or structure so that my perception of an object and my move to pick it up are fused in an intentional arc, a Gestalt or structure. I reach, for the most part, 'knowing how' to implicitly - - Polanyi's physiognosis. The body can be considered as a taken-for-granted set of performative skills, at least inasmuch as we are concerned with environmental relations. Symbolic and imaginal exercises in healing (e.g., the Simontons, 1975a, 1975b work) can in some ways be considered to be an extension of the body image in the phenomenological sense to increase 'knowing how' to control the body selfconsciously and/or to repair a break in the intentional arc at some level.

There are some interesting connections between these notions discussed and other perspectives on resonance. Firstly, the situation described as a break in the intentional arc within embodied consciousness may correspond to Marshall's posited transitional state in a complex system. Certainly, the cases studied by Merleau-Ponty are marked by an experience of betweenness, as Wild notes. Resonance, eidopoic influence, in such circumstances would amount to a transfer of 'knowing how' implicitly and with all the difficulties already mentioned regarding recovery of this information transmission. Also, if transmission is successful, then 'healing' (if this is the case) would succeed only to the extent that the eidopoic influence (healer's body image) was complete and compatible. Transplant analogies do not seem amiss here.

One of the most widely studied areas of boundary change in relation to human consciousness and action is the double-bind (Bateson, 1978; Berger, 1978). Briefly, what is involved in this process is the learning of a pattern of communication which involves fundamental contradictions regarding the meaning of experience - - of the self (psyche and body) and the other in relation. Now, an important aspect of the double-bind is that, once it is learned, then it takes only one

instance of the learned logic of contradiction to set off a whole pattern of behaviors - - most often labelled 'schizophrenic'. The double-bind victim, interestingly enough, often emphasizes one dimension of communication transmission and interpretation so that the pattern may be purely literal or metaphoric with no hint of the relevant context for understanding. The syndrome involves a 'ruptured dialectic' of a certain kind - - a splitting of communication modes and their comprising an atypical configuration. This is somewhat analogous to the hologram example: a pattern of interference is set up (split communication modes, altered boundary relations) so that one instance of contradiction evokes the whole configuration. It is a rough analogy, but one which has added implications for healing.

Just as a whole can be elicited by a part in the double-bind, some instances of healing involve changes in consciousness by the healer which are self-induced. Healing is sometimes learned in such a fashion that there is a 'betweenness' involved in the healer's social position, sense of embodiment, and relation with apparently transcendental orders (Locke, 1979). In these kinds of situations, a logic of contradiction is learned also - - one that reflects the marginality of the healer's consciousness (Meyerhoff, 1976). But, in this circumstance, as opposed to that of the double-bind victim, a change in behavior and consciousness is self-induced. A phenomenological viewpoint would suggest that self-induction might only be successful insofar as a tension is set up between the world of familiar installation of actions (everyday focal awareness) and the world(s) of potential 'otherness'.

> Shamanic balance is a particular stance - - it is a state of acute tension, the kind of tension which exists - - when two unqualified forces encounter each other, meeting head-long, and are not reconciled but held teetering on the verge of chaos, not in reason but in experience (Meyerhoff, 1976: 102).

Insofar as this tension is evident in the healer and the patient, then we have the possibility of reciprocal resonance. In that case, what distinguishes the healer from the patient will be the vectorial quality of imagery configurations and this expressed in terms of some notion of control.

As a final note, Philip Seitz (1975) gives an interesting viewpoint on resonance in psychoanalytic therapy. He discusses his own clinical cases and describes ESP-like experiences occurring between him and his patients. An important feature of his analysis is his estimation that

the ESP-like instances mostly concerned cases where he was an immature analyst - - i.e., he had not gained control over transference-countertransference dynamics; what belongs to the other as opposed to what belongs to the self. Coming to the point, Seitz explains the ESP events, arising out of poor control, as his introjection and projection of patients' anniversary reactions. In other words, a cyclical structure (memory-imagery configuration) had been learned by Seitz so that he thought of the patient and vice versa at the time of the anniversary trauma. Here, we can see elements of Levenson's configuration-learning as a basis of therapy, but without selfconscious monitoring by the healer. There are powerful resemblances between the situation of the novice psychoanalyst and the novice psychic healer who is often characterized as 'taking on' the sufferings of others before control is learned, when only a limited receptivity (boundary change) is permitted.

CONCLUSION

The purpose of this review is to look at some points of view on resonance and imagery, although these terms are not always used by the writers discussed, and to isolate the main propositions and variables relating to healing in these terms. Resonance provides a provocative angle of attack on psychic healing processes. It allows us, primarily, to consider imagery or representational systems as types of knowing - - i.e., as epistemic structures.

Out of the material presented, a definition of resonance has been constructed. It has five components. A minimal definition of resonance is covered by parts (a), (b), and (d), while the remaining two parts (c) and (e), are hypotheses concerning establishment of eidopoic influence and its execution. However, this should not be taken to indicate a definitive attitude toward the definition: it is open to review as are all the basic propositions and assumptions in the research proposal.

Resonance is a relational process wherein:
(a) there is an identity between epistemic structures in one subject and another, at the same level; and,
(b) where each level is a unity within the field of embodied consciousness, comprised of psychological and psychological-semiological elements; and,

(c) where the relation is initiated and sustained by a type of phenomenological attitude which alters the constitution (perceptual-communicative boundaries) of routine, focal awareness; and,
(d) where the epistemic structures which are in relation are imagery or representational systems which are the wholes contexting the parts of communication exchanges; and,
(e) where change in epistemic structures occurs as a function of identity and of novelty only when identity is established first, and particularly when the elements of an epistemic structure of one subject are communicated in an inverted form to that subject from the other.

REFERENCES

Ahsen, A. 'Psycheye', NY: Brandon House 1977.

Bandler, J. & Grinder, R. 'The Structure of Magic', 2 vols., Palo Alto: Science and Behavior Books, 1976.

Bateson G. 'Steps to an Ecology of Mind', NY: Ballantine, 1972.

Bateson G. 'The Birth of a Matrix or Double Bind & Epistemology', in M. Berger (ed.) 'Beyond the Double Bind', NY: Brunner/Mazel, 1978 (a).

Bateson, G. 'Breaking out of the Double Bind', Psychology Today, 12, 3, 43-51, 1978 (b).

Berger, M. (ed.) 'Beyond the Double Bind', NY: Brunner/Mazel, 1978.

Bogen, J.E. 'The Other Side of the Brain: An Appositional Mind', in R. Ornstein (ed.) 'The Nature of Human Consciousness', NY: Viking, 1974.

Broekman, J. 'Structuralism', Dordrecht, Holland: Reidel, 1974.

Castaneda, C. 'The Teachings of Don Juan: A Yaqui Way of Knowledge', NY: Ballantine, 1969.

Douglass, M. 'Social Preconditions of Enthusiasm and Orthodoxy', in R. Spencer (ed.) 'Forms of Symbolic Action', Proc. of the American Ethnological Society, 1969.

Garfield, S. & Bergin, A. (eds.) 'Handbook of Psychotherapy and Behavior Change', NY: Wiley, 1978.

Gordon, D. 'Therapeutic Metaphors', Cupertino, California: Meta Publications, 1978.

Gorney, J. 'Resonance and Subjectivity', Contemporary Psychoanalysis, 14, 2, 246-272, 1978.

Green, E. & Green, A. 'Beyond Biofeedback', San Francisco: Delacorte, 1977.

Gurman, A. 'The Patient's Perception of the Therapeutic Relationship', in A. Gurman and A. Razin (eds.) 'Effective Psychotherapy: A Handbook of Research', Oxford: Pergamon, 1977.

Gurman, A. & Razin, A. 'Effective Psychotherapy: A Handbook of Research', Oxford: Pergamon, 1977.

Horowitz, M. 'Image Formation and Cognition', NY: Appleton-Century-Crofts, 1977.

Jung, C.G. 'Synchronicity: An Acausal Connecting Principle', in W. Pauli and C.G. Jung (eds.) 'The Interpretation of Nature and the Psyche', Princeton, NJ: Princeton University Press, 1955.

Krippner, S. 'Psychic Healing - A Multidimensional View', in J. Fosshage and P. Olsen (eds.) 'Healing: Implications for Psychotherapy', NY Human Sciences Press, 1978.

Lacan, J. 'The Function of Speech and Language in Psychoanalysis', in A. Wilden (ed.) 'The Language of the Self', NY: Delta Books, 1968.

Lacan, J. 'The Four Fundamental Concepts of Psychoanalysis', London: Hogarth, 1977 (a).

Lacan, J. 'Ecrits', NY: Norton, 1977 (b).

Laing, R.D. 'The Politics of Experience', London, Penguin, 1967.

Lauer, R. 'Masters of Metaphysics', in R. Cox (ed.) 'Religious Systems and Psychotherapy', Springfield, Ill.: Thomas, 1973.

Leach, E. 'Culture and Communication', London: Cambridge University Press, 1976.

LeShan, L. 'The Medium, the Mystic, and the Physicist', NY: Viking Press, 1974.

LeShan, L. 'Toward a General Theory of Psychic Healing', in S. Dean (ed.) 'Psychiatry and Mysticism', Chicago: Nelson Hall, 1975.

Leuner, H. 'Guided Affective Imagery (GAI): A Method of Psychotherapy', American Journal of Psychotherapy, 23, 4, 4-22, 1969.

Levenson, E. 'A Holographic Model of Psychoanalytic Change', Contemporary Psychoanalysis, 12, 1-20, 1976.

Levenson, E. 'Psychoanalysis - Cure or Persuasion?', in E. Witenberg (ed.) 'Interpersonal Psychoanalysis', NY: Gardner, 1978.

Levi-Strauss, C. 'Structural Anthropology', NY: Basic Books, 1963.

Locke, R.G. 'Manipulators of Time: Symbolic Orders in Spirit Mediumship', unpublished Ph.D. dissertation, University of Western Australia, 1974.

Locke, R.G. 'Protean Principles: Explorations in the Structure & Symbolism of Healing', Opening Address, Founder's Day Conference, Psychical Research Foundation, Duke University, Durham, NC, 1979.

Macklin, J. 'Belief, Ritual, and Healing: New England Spiritualism & Mexican-American Spiritualism Compared', in I. Zaretsky and M. Leone (eds.) 'Religious Movements in Contemporary America', Princeton, NJ: Princeton University Press, 1972.

Marcuse, H. 'On Science and Phenomenology', in A. Giddens (ed.) 'Positivism and Sociology', London: Heinemann, 1974.

Marshall, N. 'ESP and Memory: A Physical Theory', British Journal for the Philosophy of Science, 10, 40, 265-286, 1960.

May, R. 'The Function of Myth in Sickness & Health', in E. Witenberg (ed.) 'Interpersonal Explorations in Psychoanalysis', NY: Basic Books, 1973.

Merleau-Ponty, M. 'Phenomenology of Perception' London: Routledge & Kegan Paul, 1962.

Merleau-Ponty, M. 'The Structure of Behavior', Boston: Beacon, 1963.

Meyerhoff, B. 'Shamanic Equilibrium: Balance and Meditation in Known and Unknown Worlds', in W. Hand (ed.) 'American Folk Medicine', Berkeley: University of California Press, 1976.

Nicolois, G. & Prigogine, I. 'Self-Organization in Non-equilibrium Systems', NY: Wiley, 1977.

Orlinsky, D. & Howard, K. 'The Therapist's Experience of Psychotherapy', in A. Gurman and A. Razin (eds.) 'Effective Psychotherapy: A Handbook of Psychotherapy & Behavior Change', NY: Wiley, 1977.

Parloff, M., Waskow, I. & Wolfe, B. 'Research on Therapist Variables in Relation to Process and Outcome', in S. Garfield and A. Bergin (eds.) 'Handbook of Psychotherapy & Behavior Change', NY: Wiley, 1978.

Polanyi, M. 'Personal Knowledge', Chicago: University of Chicago Press, 1964.

Polanyi, M. 'Knowing and Being', Chicago: University of Chicago Press, 1969.

Pribram, K. 'Languages of the Brain', Englewood Cliffs, NJ: Prentice Hall, 1971.

Pribram, K. 'Holographic Memory', Psychology Today, 12, 9, 71-84, 1979.

Prigogine, I. 'Unity of Physical Laws and Levels of Description', in M. Grene (ed.) 'Interpretations of Life and Mind', NY: Humanities Press, 1971.

Prince, R. 'Mystical Experience and the Certainty of Belonging: An Alternative to Insight and Suggestion in Psychotherapy', in E. Cox (ed.) 'Religious Systems and Psychotherapy', Springfield, Ill.: Thomas, 1973.

Rogers, C. 'The Necessary and Sufficient Conditions of Therapeutic Personality Change', Journal of Consulting Psychology, 21, 95-103,

1957.

Rogers, C. 'The Therapeutic Relationship and Its Impact', Madison, Wisc.: University of Wisconsin Press, 1967.

Rogers, C. & Truax, C. 'The Therapeutic Conditions Antecedent to Change: A Theoretical View' in C. Rogers (ed.) 'The Therapeutic Relationship and Its Impact', Madison, Wisc.: University of Wisconsin Press, 1967.

Roll, W.G. 'ESP and Memory', International Journal of Neuropsychiatry, 2, 5, 505-521, 1966.

Ryle, G. 'The Concept of Mind', Harmondsworth Middlesex: Penguin, 1976.

Seitz, P. 'ESP-like Experiences in a Psychoanalyst: A Possible Psychological Mechanism', Psychoanalysis and Contemporary Society, Vol IV, 189-209, 1976.

Shaffer, J. 'Psychefeedback', Jacksonville, Ill.: The Well-Being Center, 1978.

Simonton, C. & Simonton, S. 'Belief Systems and Management of the Emotional Aspects of Malignancy', Journal of Transpersonal Psychology, 7, 29-47, 1975.

Simonton, C. 'The Role of the Mind in Cancer Therapy', in S. Dean (ed.) 'Psychiatry and Mysticism', Chicago: Nikon Hall, 1975.

Singer, J. 'Imagery and Daydream Methods in Psychotherapy and Behavior Modification', NY: Academic Press, 1974.

Singer, J. 'The Constructive Potential of Imagery and Fantasy Processes', in E. Witenberg (ed.) 'Interpersonal Psychoanalysis', NY: Gardner, 1978.

VanKaam, A. 'Phenomenal Analysis: Exemplified by a Study of the Experience of 'Really Feeling Understood'', Journal of Individual Psychology, 15, 66-72, 1959.

Watkins, J. 'The Therapeutic Self-Developing Resonance, Key to Effective Relationships', NY: Human Sciences Press, 1978.

Wild, J. 'Foreword' to M. Merleau-Ponty 'The Structure of Behavior', Boston: Beacon, 1963.

Wilden, A. 'System and Structure', London: Tavistock, 1972.

Zinker, J. 'Creative Process in Gestalt Therapy', NY: Brunner/Mazel, 1977.

ON THE 'TRANSFERENCE' OF PSYCHIC ABILITIES: A HISTORICAL NOTE

Carlos S. Alvarado

In a recent report of psi experiments with Matthew Manning in this journal the investigators wrote about trying "to determine whether Mr. Manning could transfer his psychic abilities" to a person with which he had developed a special rapport and "who had demonstrated no such abilities previously" (Palmer, Tart & Redington, 1979, p.396).

The concept of a transference of psychic abilities or properties to an object or person that did not have that potential before is not a new idea. There are frequent mentions in the mesmeric literature of transference of animal magnetism to metals and water that then acquired healing or hypnotic properties (Dingwall, 1968). Paul Joire (1906) claimed that the human nervous force could be exteriorised and stored in substances and materials like wood, water, linen and cardboard which could then produce deflections in the sthenometer, a compass-like instrument invented by Joire to measure this force. A somewhat modified modern version of this may be the so-called psychotronic generators (Krippner, 1977).

Other observations have been made with persons. The medium D.D. Home not only showed apparent immunity to fire, but it was claimed that he could 'transfer' his ability to others (Dunraven, 1924; Home, 1921). According to Lord Adare (Dunraven, 1924), before taking a burning ember in his hand in an 1868 seance with Home, the medium made "two rapid passes" over his hand. When he picked up the ember it felt only slightly warm.

Holms (1969) presents a summary of "some remarkable manifestations of psychic force" of Mrs. Annie Abbot, who was giving public performances in the United States in 1891. She could not be moved from her physical position or stopped from falling to the floor even when several men exerted force against her. By touching a boy, she could make him just as difficult to move.

Hereward Carrington claimed to have temporarily received the psychokinetic powers of Eusapia Palladino. As he wrote: "On several occasions Eusapia transferred her power to me (seemingly) by placing her hand upon my shoulder - - - I felt nothing, but Eusapia - - - said to me, "Now, you do it!" Whereupon, if I placed my hands in front of the cabinet curtains, they would blow out, as though my hands attracted them; or I would place my hand over a small stool, and that would follow wherever I moved it - backwards and forwards - the stool moving about over the floor during the process" (Carrington, 1920, 111-112). Similar reports have been published about the 'electrostatic-PK' of Alla Vinogradova (Maigret, 1976).

These observations may suggest some kind of transference, but other variables such as suggestion (healing and hypnotic effects), a direct PK action by the medium over the target, or possible psychological psi-conducive effects such as increased motivation or decreased ownership resistance in the subject who is supposed to 'receive' the 'power', should be considered (Stanford, 1977).

REFERENCES

Carrington, H. 'Modern psychical phenomena', New York, American Universities, 1920 (Originally published, 1919).

Dingwall, E.J. (ed.) 'Abnormal hypnotic phenomena' (4 vols.), New York, Barnes & Noble, 1968.

Dunraven, Earl of. 'Experiences in spiritualism by D.D. Home', Proceedings of the Society for Psychical Research, 1924, 35, 1-285 (Originally published, 1869).

Holms, A.C. 'The facts of psychic science and philosophy', New Hyde Park, N.Y., University Books, 1969 (Originally published, 1925).

Home, Mme. D. 'D.D. Home: His life and mission' (Abridged edition), London, Kegan, Paul, Trench, Trubner, 1921 (Originally published, 1888).

Joire, P. 'The storage of the exteriorised nervous force in various bodies', Annals of Psychical Science, 1906, 4, 30-37.

Krippner, S. 'A first-hand look at psychotronic generators', in J. White (eds.) 'Future Science', Garden City, N.Y., Anchor Press, 1977.

Maigret, P. de. 'PK training in Russia', Fate, May 1976, 36-44.

Palmer, J., Tart, C.T. & Redington, D. 'Delayed PK with Matthew Manning: Preliminary indications and failure to confirm', European Journal of Parapsychology, 1979, 2, 396-407.

Stanford, R.G. 'Experimental psychokinesis: A review from diverse perspectives', in B.B. Wolman (ed.) 'Handbook of Parapsychology', New York, Van Nostrand Reinhold, 1977.

Carlos S. Alvarado,
1595 Sunnyvale 14
Walnut Creek, CA 94596
U.S.A.

CALL FOR PAPERS

Eight Annual
Southeastern Regional Parapsychological Association
Conference

The Eight Annual SERPA Conference will take place in Durham, North Carolina, on 13-14 February 1981. Anyone may submit a paper to the Program Committee, regardless of whether or not s/he is a member of SERPA or resides in the SERPA region. Original papers may be on empirical, methodological, or theoretical topic. Student papers are encouraged. SERPA members may make suggestions for topic-centered symposia.

Two copies of the papers should be sent to SERPA Program Committee, c/o Sarah Thomas Maginn, FRNM, Box 6847 College Station, Durham, NC 27708, no later than 9 January 1981. The paper must be in its final form; no incomplete papers will be accepted. Each paper should be accompanied by a 100-150 words abstract, in camera-ready form, which will be used as the basis for the abstracts published in the Journal of Parapsychology. Symposia proposals should include abstracts from all the prospective participants. Papers (but not symposia presentations) may be read in absentia with prior permission from the Program Committee. Conference information and registration materials are available from the Executive Secretary at the above address.

Robert L. Morris
Psychology Department
University of Edinburgh

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 3

NUMBER 3

NOVEMBER 1980

Robert L. Morris
Psychology Department
University of Edinburgh

EDITORIAL STAFF

Martin Johnson
Sybo A. Schouten

REGIONAL EDITORS

U.K.
Dr. John Beloff
University of Edinburgh
Department of Psychology
7 George Square
Edinburgh EH8 9JZ

Sweden
Dr. Rolf Ejvegaard
Mellansjovagen 35
S-14146 Huddinge

U.S.A.
Ms. Rhea White
2 Plane Tree Lane
Dix Hills
New York 11746

Australia
Dr. H.H.J. Keil
University of Tasmania
Box 252 C, G.P.O.
Hobart, Tasmania

W. Germany
Prof. Dr. Hans Bender
Prof. Dr. Johann Mischo
Institut fur Grenzgebiete der
Psychologie und Psychohygiene
78 Freiburg i. Br.
Eichhalde 12

Iceland
Dr. Erlendur Haraldsson
University of Iceland
Department of Psychology
Reykjavik

Italy
Dr. P. Cassoli
39, Via Valeriani
I-40134 Bologna

Prof. Dr. Emilio Servadio
Via di Villa Emiliani, 4-int.7
I-00197 Roma

The European Journal of Parapsychology publishes original experimental and theoretical papers on parapsychology. In addition, articles presenting reviews of literature, case studies and papers on related topics are accepted in so far as the subject can be considered relevant to the understanding of parapsychological phenomena or to the methodology of empirical research in parapsychology. Translations of papers originally published in a language other than English are also welcome.

Manuscripts should be written in English and should be typed, double-spaced throughout, with a wide margin, on one side of the page only. The organization of the paper must be clearly indicated by appropriate headings and subheadings. Include names of authors, positions and other affiliations, and mailing addresses. Each table

(continued inside back cover)

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

CONTENTS

Blackmore, S.J.	The extent of selective reporting of ESP ganzfeld studies	p.213
Houtkooper, J.M. Andrews, R. Ganzevles, P.C.J. Van der Sijde, P.C.	A hierarchical model for the observational theories: A study of subject identity and feedback in repeated retroactive psychokinesis	p.221
Harley, T.A. Sargent, C.L.	Can experimenters localize nonrandom target sequences in random number tables?	p.247
Schouten, S.A.	Emotions as targets in ESP studies	p.253
Shafer, M.G.	Exploratory observations with PK metal benders from Southern California	p.273
Keil, H.H.J.	The voice on tape phenomena: Limitations and possibilities	p.287
Lucadou, W.v. Kornwachs, K.	Development of the system theoretic approach to psychokinesis	p.297

PUBLICATION POLICY

Twice a year the Parapsychology Laboratory of the University of Utrecht publishes the European Journal of Parapsychology. The object of the European Journal of Parapsychology is to stimulate and enhance the activity in this field, especially in our corner of the world, by communicating research results and issues related to professional parapsychology. Although there will be an emphasis on experimental work, theoretical articles are also welcome. Contributions from all over the world will appear in the journal.

A hallmark of the European Journal of Parapsychology is the attempt to avoid selective reporting, that is, the tendency to bury 'negative' results and only to publish studies that 'turn out'. To avoid turning the journal into a graveyard for all 'unsuccessful' studies, we require that the acceptance or rejection of a manuscript should take place prior to the phase when the experimental data are collected. The quality of the design and methodology and the rationale of the study are considered more important than the level of significance of the outcome of the study. As a practical rule, we advise the potential contributor of an article to submit the design of his planned study before the work is actually carried out. The rationale of the study should be stated, as well as all the hypotheses related to it. Furthermore one should try to specify the number of subjects, the number of trials, etc., plus the type of statistical methods one plans to use for evaluation.

Priority will be given to the publication of studies which fulfil the above-stated publication policy.

The final manuscript with presentation of results must reach us two months in advance of the official publication dates, which are May 1st and November 1st.

THE EXTENT OF SELECTIVE REPORTING OF ESP GANZFELD STUDIES

Susan Blackmore
Parapsychology Laboratory
University of Utrecht

The repeatable psi experiment has proved an elusive goal for parapsychology. However, certain recent methods for eliciting psi, notably the use of ganzfeld stimulation, seem to have provided some degree of replicability. In 1977 Honorton reviewed all ganzfeld studies then published and showed that half, eight out of sixteen, obtained significant results. In a second review (Honorton, 1978) he gave 14 out of 26, or 54%, as significant. Sargent (1979) suggested that the ganzfeld technique has given 58% replicability both across laboratories and studies. My own survey of the literature published to date shows that of 31 studies 18, or 58%, obtained significant results. These figures are shown in table 1.

These estimates appear encouraging. However, various factors have to be taken into consideration before accepting them as valid. These include the methodological weaknesses in many of the studies, the use of inappropriate statistics or overanalysis, and selective reporting. Sargent (1980a) separated studies into those he considered adequate and inadequate in methodology. Of 11 deemed adequate 6, or 54%, obtained significance. In my own survey I considered 12 published studies to be adequate. Others were ruled out for such reasons as the

---

Note: This paper fulfils the publication policy of this journal.

TABLE 1
Estimates of the repeatability of ESP in the ganzfeld (published studies)

	Total	Number of significant studies	Percentage of significant studies
Honorton 1977	16	8	50%
Honorton 1978	26	14	54%
Sargent 1979	24	14	58%
Blackmore 1980	31	18	58%

use of picture targets without a duplicate set for judging, or the possibility of communication between subject and agent prior to judging. Some studies can be considered adequate for results obtained from independent judging but not from subjects' judging. Of the 12 remaining 6, or 50%, were significant. Clearly the adequacy or otherwise of the methods used makes little difference to the apparent replicability of the technique.

Selective reporting may also be a relevant factor. If many experiments are carried out but those which seem to be giving chance results are abandoned before completion or are never written up for publication then there will be a bias towards significance in the literature. This bias may be introduced by subjects dropping out, by authors themselves, or by editors selecting submitted work for publication. Any estimates of the repeatability of a finding based only on published results may overestimate the true repeatability to an extent dependent on the size of that selection. Sargent has pointed out that a ganzfeld study is relatively time-consuming and arduous to perform as compared with many types of forced-choice psi experiments. They are therefore unlikely to be undertaken lightly, for example as student projects or informal experiments. In addition, since the first such study was reported only six years ago (Honorton and Harper, 1974) there has not been a long time in which large numbers of such studies

could have been carried out. One might therefore expect the number of unreported ganzfeld studies to be small, and certainly smaller than for many other types of psi experiment. Finally the method used varies comparatively little.

For all these reasons it seemed that it might be possible to find out about the majority of unpublished ganzfeld studies and to compare the success rate of those with that of the published ones. In this way an estimate of the extent and importance of selective reporting could be made. The published data was surveyed, as already noted, and a questionnaire was used to find out about unpublished work. The results for each were then compared.

METHOD

A questionnaire was sent to any person who it was thought might have been associated with ganzfeld studies. This asked about studies not completed, and those completed but not published. The complete questionnaire appears in the appendix.

The questionnaire was initially sent to 81 people, most of whose names and addresses were obtained from the Parapsychological Association member's list. For those who had published ganzfeld work a list of references was appended to check that none had been omitted. In response to information given by some respondents a further five questionnaires were sent out, making 86 in total.

RESULTS

40 completed questionnaires were returned, giving a return rate of 47%. These elicited a total of 32 further ganzfeld studies, reported by 15 respondents. Of these 12 studies were reported as not completed. These included seven "in progress" with the intention to complete and five others. The other reasons for not completing were "methodological defects" (2), "results not significant" (2), "demonstration to the media, aborted" (1), "high attrition rate and overly complex design-analysis requirement" (1), and "departed for graduate school" (1) (the total is more than five since respondents could give more than one reason for not completing the study). In no case was "results

not significant" given as the sole reason for failing to complete and therefore no selection at this stage was apparent (assuming thruthful answers).

20 completed but unpublished studies were listed. Reasons for not publishing were as follows:

Intending to publish	13
Methodological defects	5
Results not significant	1
Other and:	
Never intended for publication	2
Awaiting modified confirmation	1

The 20 unpublished studies included both significant and non-significant results (significance on primary analysis for ESP in the ganzfeld only was asked for). The results of one could not be evaluated for various reasons given by the author. Of the remaining 19, 7 were reported to have given results significant at the .05 level on the primary analysis. This is 37% which compares with 58% (by my estimate) in the published data. As might be expected, the difference is in the direction of more significance in the published work, but a chi-square test comparing significant and non-significant for published and unpublished shows no significant difference (see table 2).

The adequacy of the method used is also relevant. Authors may be more willing to report significant findings if the method is sound, and the method may affect the findings. For any analysis here it is necessary to assume that authors correctly reported whether their methods were adequate or not and this assumption may be unwarranted and is hard to test. However, assuming so, we find that 14 of the 19 studies were adequate and of these 5, or 36%, obtained significant results. This provides no evidence that methodological weaknesses were responsible for significant findings in the unpublished any more than in the published work.

DISCUSSION

The proportion of 'significant' overall psi scores for unpublished ganzfeld studies is comparable to that found in the published ones. We

TABLE 2
Proportions of 'significant' studies,
published and unpublished

	Total	Number of significant studies	Percentage of significant studies
Published	31	18	58%
Unpublished	19	7	37%
Total	50	25	50%

cannot be sure that all unpublished have been detected here, but certainly all those from the major parapsychological laboratories have been (assuming honest replies), and there would seem to be no a priori reason to suppose that the proportion would be markedly different in those which were undetected by this survey. In fact, in order to make a difference to the conclusions reached here we would need to find a further four non-significant unpublished but completed studies, and no significant ones, to add to those included here. It seems unlikely, though certainly possible, that with further investigation the position would change in this way.

Also relevant is to consider the overall significance. Two years ago Honorton (1978, p.87) pointed out that "even if we make the absurd assumption that there are ten . . (unreported failures)..for each significant study, the observed results would still be significant at p=0.02". Of all the studies included here, both published and unpublished, 25 out of 50 provided results significant at the .05 level. If these 25 were to be the selected best of a larger number of studies there would need to be a further 450 undetected non-significant studies.

We may therefore conclude with some certainty that the bias introduced by selective reporting of ESP ganzfeld studies is not a major contributor to the overall proportion of significant results,

and the apparent success of the technique. If the positive results are not due to psi but are spurious this must be for some reason such as methodological errors or overanalysis undetected here, unconscious biases, untruthful answers, or deliberate fraud, or other such factors. It is not due solely to selective reporting.

ACKNOWLEDGMENT

I would like to thank all those respondents who completed and returned the questionnaire so promptly.

ABSTRACT

Estimates of the replicability of ESP in the ganzfeld based on the published data alone range from 50% to 58%. In a survey of unpublished ganzfeld studies a further 32 studies were elicited. Of 20 completed but not published 37% obtained significant results in the primary analysis, a proportion not significantly different from that in the published data. It was concluded that selective reporting is not a major contributor to the apparent success and repeatability of ESP in the ganzfeld.

REFERENCES

Honorton, C. 'Psi and internal attention states'. In Wolman, B. (ed.) 'Handbook of parapsychology', Van Nostrand Reinhold, 1977.

Honorton, C. 'Psi and internal attention states: Information retrieval in the ganzfeld'. In 'Psi and states of awareness', Parapsychology Foundation, 1978.

Honorton, C. & Harper, S. 'Psi mediated imagery and ideation in an experimental procedure for regulating perceptual input'. J. A.S.P.R., 68, 156-168, 1974.

Sargent, C.L. 'Repeatable significance and the significance of repeatability'. Paper presented at the 3rd International Conference of the Society for Psychical Research, Edinburgh, 1979.

Sargent, C.L. Personal communication, March 1980 (a).

Sargent, C.L. Personal communication, October 1980 (b).

Present address:
Pear Tree Cottage
Greyfield Road
High Littleton
Avon, BS18 5YB
England

APPENDIX

State University of Utrecht

Parapsychology laboratory

Varkenmarkt 2
Utrecht 2501
The Netherlands
Telephone 030-331211

Dear Colleague,

There has been much discussion about the replicability of ESP in the ganzfeld. In his review of the literature Honorton (1977) has pointed out that the results may be contaminated by reporting bias, and many studies suffer from methodological weakness. For these reasons it is important that a complete list of all ganzfeld studies be compiled. This should be possible since all this work is of relatively recent date. I would therefore be extremely grateful if you would assist me by completing the enclosed questionnaire. If you have published ganzfeld studies a list of them is enclosed. If there are others others I have not included I would be greatly obliged if you could reference them. If you have students or colleagues who have carried out unpublished ganzfeld studies perhaps you could also ask them to complete a copy of the questionnaire.

With many thanks in anticipation,

Susan Blackmore

QUESTIONNAIRE

Please answer the questions separately for each relevant study. You may need to copy the questionnaire or give further answers on a separate sheet.

1. How many ganzfeld studies have you started but not completed ? ____

2. Why did you not complete them (if any) ? (tick one for each study)

Intending to complete	____	____	____	____
Methodological defects	____	____	____	____
Results not significant	____	____	____	____
Other	____	____	____	____

3. How many ganzfeld studies have you completed but not published ? ____

4. How many of these produced overall significant psi-hitting (by sum of ranks, z, binomial p) ? ____

5. Why did you not publish ? (tick one for each study)

Intending to publish	____	____	____	____
Methodological defects	____	____	____	____
Results not significant	____	____	____	____
Other	____	____	____	____

6. Would you be willing to send further details and complete data if requested ? Yes/No

7. Do you know of anyone else who has completed, but not published any ganzfeld studies ? _____

8. Any further comments.

A HIERARCHICAL MODEL FOR THE OBSERVATIONAL THEORIES: A STUDY OF SUBJECT IDENTITY AND FEEDBACK IN REPEATED RETROACTIVE PSYCHOKINESIS.

Joop M. Houtkooper
University of Amsterdam

Robin Andrews
Paul G.J. Ganzevles
Pieter C. van der Sijde
Research Institute for
Psi-phenomena and Physics

The purpose of this study is to explore the roles of 'consciousness' and feedback within the observational theories (Millar, 1978).

The central axiom of observational theory (OT) is that the act of observation of an event with an uncertain outcome influences that outcome. This is not falsifiable in principle because there can be no evaluation of an outcome without observing it. Although it may seem that observational theory is thus in much the same position as, e.g., a theory of the hereafter, it differs in that OT has lent itself to the construction of different specific models, like the multiplicative model (Schmidt, 1975, 1978b) and the additive-information model (Walker, 1975, 1977). Within these a productive effort can be spent on

---

Note: This paper fulfils the publication policy of this journal

the falsifiability of OT. The divergence problem in particular has attracted considerable interest (Bierman and Houtkooper, 1980, Millar, 1978, Hartwell, 1977, Millar and Hartwell, 1979, Houtkooper, 1977b, Schmidt, 1975, 1978a, 1978b, Walker, 1977), discussed later.

However, our present approach is to regard OT as just a framework within which various models of parapsychological phenomena can be constructed. We propose here a new model in which it is assumed that both the random number generator (RNG) and the previous observers act as absorbers of psi.

The multiplicative model can be developed as a mathematical formalism hinging on feedback without ever introducing the notion of observation. However Walker has developed this explicitly in his theory together with, or from, his ideas about consciousness, which he treats almost like a physical variable (Walker, 1970).

In this paper we take up this wide-ranging concept at a more common-sense level, that consciousness is attached to individuals and that it has continuity in time. In this sense the characteristics of consciousness seem identifiable with those of memory or individuality and although not perfect, we will use this latter term here. The role of individuality in connection with repeated retroactive psychokinesis (RRPK) will be elaborated in the next section.

The other aspect of this study, feedback, is manipulated firstly as the length of the feedback time interval and secondly by giving no feedback at all about the pre-recorded trials in what would otherwise be a RRPK experiment.

The notion that the results of parapsychological experiments may be due to the experimenters and not, or much less so, due to the subjects is taken seriously in this study. The evaluation of the outcome of the experiment by the analyzers certainly constitutes a form of feedback and this process is researched in the present study, thereby extending the part of the experimental procedure that is under scientific scrutiny.

The experiments described here were carried out as a student research project of the "International Summerschool Experimental Parapsychology 1980", organized by the Research Institute for Psi-phenomena and Physics (RIPP).

THEORY

In the version of OT we call the multiplicative model, a psi-influence is considered to change the ratio of hits to misses of a RNG by a factor (the psi-strength) which is only dependent on the subject, or more abstract, the psi-source (Schmidt, 1975). As the psi-strength is assumed to be invariant under different conditions, like delayed feedback (retroactive psychokinesis or retro-PK), complexity of the RNG and influencing RNG events which have been previously observed, a simple structure emerges with elegant mathematical properties. The only drawback is the divergence problem which can be stated in two ways (Millar, 1978): First, the many observers paradox, or the combined influence of a large number of observers producing an extreme result (Houtkooper, 1977b). Secondly, there is the experimental control dilemma, since no experiment is ever finished in principle.

These problems stem from the incorporation of time-independence. However, time independence has been one of the fruits of the observational tree that has made a big impact. Although we are still stunned with the impression that time-independence is one of the hallmarks of the ideal theory of psi (but no doubt also of good science-fiction), the trouble is that the omni-explainability it provides is only the guise of the non-falsifiability that it implies. Time independence should therefore elicit scrutiny instead of devotion. Here we split it into two possible components: First, the assumption of independency of the time interval elapsed before an observation and secondly, the assumption of independency of the order in which different observations take place. This last assumption we call into question, thereby opening up the possibility of new models of psi-phenomena. The same line of reasoning has been followed by Bierman (1980) in his discussion of the apparent failure of the 'potential observer theory' (see also Bierman and Houtkooper, 1980).

Here we introduce the hierarchical model, in which a later observer of a random event is considered to be observing and acting on the system of preceding observer plus the system he observes. Both the preceding observer and the system observed by him, which consists of still earlier observers plus the RNG, are considered as 'absorbers' of the psi-influence of the later observer. This assumption is quite plausible because the preceding observer is also determining the result and he has also to be 'moved' so to speak.

Bierman's (1980) argument on this point is that the system of observer plus an observed quantum-mechanical random system is again a quantum-mechanical random system. This would on itself rather argue for a susceptibility to an influence of such a system equal to the susceptibility of a simple RNG, instead of a diminishing influence of later observers.

Our argument is that the preceding observer is a co-absorber of the psi-influence of the later observer. If this influence is split equally (or in another constant ratio) between the preceding observer and the system that is one level less complex, we get a geometrically decreasing influence with the number of earlier observers and therefore a convergent effect.

This model has mathematical properties which can be compared with the properties of the model proposed by Hartwell (1977) who postulated an influence of later observers proportional to the inverse square of the observer number. This is another way to obtain a mathematically convergent result, but the present model has the advantage of being structurally formulated, instead of purely mathematically. This means that details of the structure rather than variations of a formula can be tested experimentally, which gives in our opinion a clearer insight than when one just tries a few out of a multitude of formulae. Although the differentiation between structures and formulae is at one level purely arbitrary, it is not when one limits them both (structures and formulae) by the criteria that they have to make sense, be elegant and fruitful, or more strictly, that they have to contend with Frege's principle of compositionality (Frege, 1923).

In the present study we try to get at the detailed structure by considering repeated observations of the same random events, on the one hand by different subjects and on the other hand by the same subject each time.

In the hierarchical model there is no reason why the repeated observation of one set of pre-recorded trials (PRTs) by different subjects should give the same result as repeated observations by the same subjects (under the assumption of equal psi-strengths). However, this difference depends on the nature of the PK interaction of which we give two aspects:

First, if we regard PK as the back-action of the information transfer from the random event to the observer (Walker, 1975,1977,

Schmidt, 1975), repeated observations will only have an effect when there is information to transfer. If 'information' is used in the sense of 'what annihilates uncertainty', transfer of information can only occur in the case of different observers (without information transfer between them!) or in the case where the observer has completely forgotten about his earlier observation. Thus the memory of the repetitive observer is important, but nevertheless we predict from this conjecture that the scoring enhancement by repeated observation (Schmidt, 1975) will be less in the case where the same observer acts as compared with multiple observers of the same events.

Secondly, a line of reasoning can be made, independent from the first in principle, that gives a different prediction. For the sake of clarity of our argument however, we forego the strict definition of information transfer made above. With a more loose definition of information, where e.g. identical information is also information, we can look at the assumption of the hierarchical model that the preceding observer acts as an 'absorber' of psi together with the one-level less complex random event generating system. If the actual observer is the same individual as the preceding observer, there is the possibility that he doesn't have to 'move' the preceding observer, i.e. that he doesn't absorb his own psi. In other words, he has a special information channel to himself as the preceding observer, consisting of the identity of himself as an individual, or the continuity of his individual consciousness. If these concepts carry too much surplus-meaning, consider his memory as the information channel from one observation to the next. The result of this information channel is here an enhanced effect, which is in the 'individual subject' case equal to the effect in the multiplicative model, as compared with the 'multiple subject' case, where the result of the purely hierarchical model holds. The enhancement due to 'individuality' as an information channel is in principle not even limited by the result of multiplicative model, namely a linear increase with the number of observations. For instance, the vividness of the memory-trace of the observed events might play a role, so that we get on the borderline between abstract modelling and psychological effects.

Summing up, the two lines of reasoning we have given predict opposite results for what we call individuality: The difference between the individual subject and multiple subjects performing repeated observations. The question common to both arguments is whether an individual absorbs his own psi (from the future) more

readily, or less readily than the psi from someone else. If there is a difference, something might be learned about the psi-process. Moreover, it is not at all inconceivable that what in reality occurs, depends on the memory characteristics of the subject, and thus is different for different subjects. Our conclusion is, to clarify the issue of individuality by experiment. As the concept of information plays a crucial role here, it is a methodologically good choice to implement types of feedback with different information content. For a discussion of feedback in OT, see Millar (1978).

THE PILOT EXPERIMENT

The experimental setup consists of an electronic random number generator (RNG) generating random events with two equiprobable outcomes, which is connected to a computer to register results and to give visual feedback to the subject. For a more detailed description see Bierman and Houtkooper (1975). The feedback for each experimental session consists of a series of runs, each divided in time-intervals. For each interval, on the display screen a horizontal deflection of a line is projected which is proportional to the z-score of the difference between the numbers of trials on both channels of the RNG, taken cumulatively from the start of the run. For the trials of the RNG, the computer can be programmed to take one of the following two cases:

1. Momentaneously generated trials only (MGTs), which means that the aim is straight real time PK.
2. Momentaneously generated trials (MMTs) mixed with pre-recorded trials (PRTs). In other words, real time PK mixed with retro-PK.

In both cases the parameters of the series of runs (e.g. see table 1A) are obtained by the computer from the 'input data file', which furthermore may contain pre-recorded trials. The momentaneously generated trials, either MGTs or MMTs, are registered during the experiment in an 'output data file' of similar form as the 'input data file'.

In the first case these are ignored and the feedback to the subject shows the result of only the trials generated at that time by the RNG (straight real time PK).

In the second case the PRTs are added to the MMTs, per interval and per RNG channel, and on the result of this addition a z-score is calculated for display to the subject. In this case the feedback contains twice the number of trials of the first case, but as the result is displayed as a z-score, it is not possible to distinguish the cases visually. The reason of mixing PRTs with MMTs is to avoid to give exact duplicate feedback in repeated retro-PK, which might have unwanted psychological effects on the subject.

In some previous experiments with RRPK both possibilities, mixed PK and straight real time PK, were alternated (Houtkooper, 1977a), whereas in the present experiment we planned to have only the mixed PK case. The pilot experiment was planned to demonstrate the possibility to test the individuality hypothesis as discussed in the previous section. Each of the 4 subjects, who were the authors themselves, performs 8 sessions (see table 1A).
Each session consists of a series of 4 runs with interval lengths of 15 and 0.5 s, with target directions alternated as ABBA. The frequency of the RNG is for both experiments 30 Hz. The assignment of the subjects to the 8 PRT series or sets of data, each of which is observed 4 times, is given in table 1B.

The order of the sessions is determined by a random choice whether e.g. subject A first observes PRT series no. 1 or no. 5 (see table 1B). These random choices were generated before the start of the experiment by the first author, who had to instruct the computer before each session as to the PRT series involved. Therefore, he was aware of the condition of individuality for the sessions he performed as a subject (and also for the other sessions). The other 3 subjects were not informed about the individuality condition however. This imperfection was regarded as not a serious problem in view of the mainly demonstrative purpose of the pilot experiment. After the completion of the 32 sessions, the results of the experiment were analyzed with 3 of the authors present. Beforehand we decided to focus in the analysis on the PRTs, without putting forward any formal hypothesis. Of most interest was the individuality effect, which is predicted to be present only in the PRTs. For the PRTs we give some raw data in table 2. An important note for the interpretation of these data has to be made at this point: The PRTs in the pilot experiment were inadvertently not observed at all during the formal experiment!

However, we did not realize at the time of analysis this had happened, due to one erroneous instruction to the computer. This error

TABLE 1A
Conditions of the session in the pilot experiment

Run	Target direction	No. intervals	Length of interval
1	left	30	0.5 s
2	right	30	0.5 s
3	right	1	15 s
4	left	1	15 s

TABLE 1B
Design of the pilot experiment; the subjects A,B,C,D assigned to the 8 PRT series.

Condition	Individual subject				Multiple subjects			
PRT series	1	2	3	4	5	6	7	8
Obs. order								
1	A	B	C	D	A	B	C	D
2	A	B	C	D	B	A	D	C
3	A	B	C	D	C	D	A	B
4	A	B	C	D	D	C	B	A

caused the computer to carry through the experiment with the display of MGTs only (according to the first case mentioned above). However, the overall result of the unobserved PRTs is highly significant psi-hitting with z=+3.23; p=.0013 two-tailed! Our initial enthusiasm

TABLE 2
PRT's of the pilot experiment, split for individuality.

Condition	Target direction	Hits minus misses	Number of trials
Individual subject (IS)	left	+102	3600
	right	+95	3599
	total	+197 (z=2.32)	7199
Multiple subjects (MS)	left	+101	3599
	right	+90	3600
	total	+191 (z=2.25)	7199
Overall		+388*	14398

*) z=3.225, p=.0013

for this finding changed into astonishment when we checked on the MGTs and found out that only these had been in the feedback. From table 2 it appears that the individuality effect, which is now rather meaningless, is very small: z = +0.050, with the individual subject (IS) condition showing slightly higher scoring than the multiple subject (MS) condition. Further analysis gives a nonsignificant time interval effect: z=-0.868, with 15 s intervals scoring more than half second intervals. The results per subject are given in table 3 as z-scores.

In this table the PRTs in the MS condition are omitted, as these are associated with every subject.

We see the main effect in the IS condition (see also table 2) being

TABLE 3
Z-scores per subject for main and differential effects

Subject / CONDITION effect	A	B	C	D	Overall
IS: PRT's					
Main	+2.22*	+0.09	+2.36*	-0.03	+2.32*
Interval	-0.33	+0.94	-0.85	-1.53	-0.88
ALL MGT's					
Main	-0.04	-1.05	-0.28	+0.63	-0.37
Individuality	+0.17	-2.03*	-0.71	+0.41	-1.08
Interval	+0.02	-0.83	+0.49	-1.32	-0.82
IS: MGT's					
Main	+0.09	-2.18*	-0.70	+0.73	-1.03
Interval	-1.13	-1.28	-0.11	-2.00*	-2.26*

$p<.05$, two-tailed

Note: Interval effects are differences, taken to be positive if 0.5 s intervals score higher, or less negative, than 15 s intervals. The individuality effect is given as the z-score of the difference of IS minus MS conditions.

contributed almost entirely by Ss A and C. A chi-square test on the variance between subjects is not significant. Furthermore, in the MGTs in the IS condition, which were analyzed for comparison with the result of the PRTs, there is a significant difference between the short and the long time interval conditions, the short intervals scoring significantly negative ($z=-2.33$), while the long intervals score only slightly positive (i.e. in the target direction). The main conclusion from the pilot experiment is that a most puzzling effect is

obtained in the PRTs. In the framework of OT it can only be ascribed to the analyzers of the results as the first observers of the data. The accidental character of this experimental condition makes it hardly possible to replicate this finding.

THE SECOND EXPERIMENT

For the reasons mentioned above, the second experiment cannot be a confirmation of the pilot. Rather, the pilot experiment stimulated us to develop a better experiment. First of all, the randomization of the experimental conditions over the sessions was automated, so that none of the subjects (the authors) was aware of the individuality condition and of the newly introduced feedback condition of the PRTs. The finding of an effect in non-observed PRTs in the pilot experiment suggested that we manipulate this feedback condition, so that in half of the sessions there was mixed PK and in the other half straight real time PK. When there is straight real time PK, the sets of PRTs only associated with this condition are not observed. These non-observed or hidden PRTs as we will call them are associated with the session through the run-parameters being obtained from the 'input data file' that also contains the PRTs. Of course, this is a rather tenuous association.

A second modification suggested by the pilot is that, analyzer effects possibly playing a role, an analyzer split is designed so that each analyzer got equal parts of data of each subject in each condition, the smallest undivided section of data being the runpair (two adjacent runs). The task of the analyzers is to evaluate the outcome for the first four of the formal hypotheses, based on his own part of the data in the form of a summary table produced by the computer.

A third modification is made to be able to discern the time-interval effect from a within-session decline effect. In the pilot experiment this distinction cannot be made, because long time intervals occur always in the second half of the session (see table 1A). In the second experiment a session consists of 8 runs with the target directions of the runs alternating as ABBA and with interval lengths of 30 s and 0.25 s alternating as ABBA over the runpairs, see table 4A. The assignment of subjects to the 4 subsequent occurrences of each of the 16 PRT series is given in table 4B. The randomization of the

conditions for each subject is done automatically by the computer, so that the subjects are blind for the feedback condition (observed vs. hidden PRTs) and the individuality condition.

HYPOTHESES OF THE SECOND EXPERIMENT

The following hypotheses were formulated for the second experiment:

H1: In the observed PRTs there will be an overall deviation from chance expectation.
H2: There will be a difference in scoring between observed and hidden PRTs.
H3: In the observed PRTs there will be a difference in scoring between the individual subject (IS) condition and the multiple subjects (MS) condition.
H4: In the observed PRTs there will be a difference in scoring between the time interval conditions: Feedback over a run in one 30 s interval vs. 120 intervals of a quarter of a second.
H5: For the overall effect in the observed PRTs there will be an increased variance between analyzers, who analyze parts of the data resulting from a random split.

The hypotheses are to be evaluated: For H1 as the z-score of the deviation from chance expectation; for H2, H3 and H4 as the z-score of the difference; for H5 as the chi-square value of the variance (obtained by muliplying the variance by the number of degrees of freedom).

The statistical tests are two-tailed for H1, H2, H3 and H4 and one-tailed for H5, all at the 5 percent level of significance.

The random split of the data for the analyzers is performed as follows: Each of the analyzers is assigned randomly one of the 4 runpairs of PRT series 1. The following PRT series are split by taking for each analyzer the succeeding runpair of the succeeding PRT series. E.g. if analyzer A is assigned runpair 3 of PRT series 1, he will get runpair 4 of series 2, runpair 1 of series 3, etc.

We note that the first two hypotheses are not independent, as H2 tests the difference between the result for H1 (the observed PRTs) and

TABLE 4A
Conditions of the session in the second experiment

Run	Target direction	No. intervals	Length of interval
1	left	120	0.25 s
2	right	120	0.25 s
3	right	1	30 s
4	left	1	30 s
5	left	1	30 s
6	right	1	30 s
7	right	120	0.25 s
8	left	120	0.25 s

TABLE 4B
Design of the second experiment: Subjects A,B,C,D assigned to the 16 PRT series.

Feedback condition	PRT's observed								PRT's hidden							
Individuality condition	MS				IS				MS				IS			
PRT series	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Observation order																
1	A	C	B	D	D	B	C	A	B	D	A	C	D	B	C	A
2	C	A	D	B	D	B	C	A	D	B	C	A	D	B	C	A
3	B	D	A	C	D	B	C	A	A	C	B	D	D	B	C	A
4	D	B	C	A	D	B	C	A	C	A	D	B	D	B	C	A

the result for the hidden PRTs. This choice of hypotheses was caused by the ambiguous picture of the world we had and the reader may also have at this point, namely that on the one hand only the observed PRTs, according to the central axiom of OT, should contain an effect (see H1) but on the other hand the strong effect in the hidden PRTs seems to disprove this. We will return to this point in the discussion.

RESULTS OF THE SECOND EXPERIMENT

The data of the second experiment are given in table 5 for the observed and the hidden PRTs, split for individuality. Results for the PRTs per analyzer are given in table 6.

In both tables the results for hypotheses H1, H2 and H3 can be read, while the results for H4 and H5 are given in table 6 only.

Summarizing:

H1: The overall effect in the observed PRTs: z=+0.371; n.s. H2: The difference between observed and hidden PRTs: z=+1.995; p=.046.
H3: The difference between short and long intervals in the observed PRTs: z=-1.354; n.s.
H4: The difference between the IS and MS conditions in the observed PRTs: z=-0.571; n.s.
H5: The variance between the analyzers for the main effect in the observed PRTs: Chi-square=0.680, with 3 degrees of freedom; n.s.

Therefore, of the formal hypotheses only H2, on the difference between observed and hidden PRTs is significant. For the largest part, this difference stems from the effect in the hidden PRTs. Also in the hidden PRTs, the between-analyzer variance of the main effect is significantly larger than chance expectation, while this variance is a priori independent from the overall (main) effect. The MGTs and MMTs are analyzed only exploratively (see table 7, which is the companion table of table 6 for the PRT results). In table 7 we note that significant psi-missing (z=-2.20; p=.028) occurs in the MMTs, which have been observed together with the observed PRTs, see H1. The between-analyzer variance of this result is marginally significant

TABLE 5
PRT's of the second experiment split for individuality and feedback condition

Feedback condition for PRT's	Individuality condition	Target direction	Hits minus misses	Number of trials
Observed	IS	left	-93	14399
		right	+69	14399
		total	-24	28798
	MS	left	+95	14399
		right	+18	14398
		total	+113	28797
	TOTAL		+89	57595
Hidden	IS	left	-83	14397
		right	-155	14397
		total	-238	28794
	MS	left	-230	14398
		right	-120	14398
		total	-350	28796
	TOTAL		-588*	57590

*) z=-2.45, p=.014, two-tailed

TABLE 6
PRT's of the second eexperiment split for analyzers, given as z-scores.

Analyzer CONDITION effect	A	B	C	D	Overall	Chi-square between analyzers
OBSERVED PRT's					H1:	H5:
Main	+0.22	+0.49	-0.50	+0.52	+0.37	0.68
					H3:	
Individuality	-0.84	+0.18	+0.30	-0.78	-0.57	1.10
					H4:	
Interval	-0.32	-2.72*	+0.53	-0.19	-1.35	6.02
HIDDEN PRT's						
Main	-1.45	+0.65	-3.64*	-0.46	-2.45*	10.00*
Individuality	-0.38	-0.67	+0.48	+1.51	+0.47	2.87
Interval	-1.52	+0.38	-0.49	+0.12	-0.75	2.14
ALL PRT's						
Main	-0.86	+0.80	-2.92*	+0.04	-1.47	7.76
					H2:	
Feedback	+1.18	-0.11	+2.22*	+0.70	+1.99	2.86

*) $p<.05$, two-tailed

Note: The formal hypotheses are indicated by H1 to H5. The variance between analyzers is given as a chi-square value with 3 degrees of freedom. The individuality effect is given as the difference of IS minus MS conditions. The interval effect is given as the difference of the 0.25 s minus the 30 s condition. The feedback effect is the difference of observed minus hidden PRT's.

(chi-square=7.77; p=.051). This variance in the MMTs corresponds with H5. Furthermore, the significant difference between MGTs and MMTs (z=-2.02; p=.044), although not independent from this psi-missing in the MMTs, corresponds with H2. It is associated with a significantly increased variance between analyzers (chi-square=8.44; p=.038). Because the subjects are the same persons as the analyzers, we give the results split for subjects in table 8.

This table can also be compared with table 3 with the results of the pilot split for subjects.

DISCUSSION

The two experiments in this study have come up with some findings that do not fit easily in the theoretical picture we sketched. In the hidden PRTs of the pilot we have a significant effect that at first sight seems to contradict the central axiom of OT, as these trials were not observed at all by the subjects. The analysis procedure however involves looking at the results and therefore is a possible source of influencing them. No conceivable experiment of which the outcome would become known can escape this fate, so that the central axiom of OT cannot be disproven in this way. In other words, the central axiom is non-falsifiable. Although this is bad (Millar, 1978), before we throw away our explanation because of this reasonable criticism, we should look whether we can find any corroborative evidence for an analysis effect. One way to do this is to look for an analyzer effect, that is, the difference in analysis effects of different analyzers. Another way is to look for an explanation why these trials are particularly susceptible to analysis effects.

The first approach is done by looking at the hidden PRTs of the second experiment, as the pilot had no analyzer split. As we have seen, here we are so lucky to get a significant effect, which would have been a confirmation if we had been so bold to hypothesize it! That the effect changes sign between pilot and second experiment would have been a little troublesome, but this is usually accounted for by two-tailed testing of the hypothesis. The problem we had choosing our hypotheses is already mentioned (Lucky enough for our consciences we have preregistered the experiment with this Journal's editor, so that hypothesis massage is out of the question). We have chosen to stick to our theoretical framework as we saw it, but without ignoring

TABLE 7
The momentaneously generated trials, MGT's and MMT's of the second experiment, split for analyzers, given as z-scores.

Analyzer CONDITION Effect	A	B	C	D	Overall	Chi-square between analyzers
MMT's						
Main	+0.23	+0.35	-2.43*	-2.55*	-2.20*	7.77
Individuality	-1.32	+1.53	+1.81	-0.85	+0.58	7.77
Interval	-0.30	+0.10	+0.55	-0.50	-0.08	0.65
MGT's						
Main	-1.18	+0.72	+1.08	+0.70	+0.66	3.15
Individuality	+0.01	-0.56	-0.45	-0.59	-0.79	0.23
Interval	+0.84	-1.48	-0.38	+0.84	+0.19	4.94
MMT's and MGT's						
Main	-0.67	+0.76	-0.95	-1.31	-1.08	2.47
Feedback	+1.00	-0.26	-2.48*	-2.30*	-2.02*	8.44*

*) $p<.05$, two-tailed

Note: The feedback effect is given as the difference of MMT's (mixed PK) minus MGT's (straight real time PK). See also note with table 6.

TABLE 8
Results of the second experiment, split for subjects, given as z-scores.

CONDITION / Effect	A	B	C	D	Overall	Chi-square between subjects
OBSERVED PRT's, IS						
Main	-0.99	+1.51	+0.91	-1.71	-0.14	6.98
Interval	-0.47	-1.45	+0.55	-1.02	-1.20	2.25
HIDDEN PRT's, IS						
Main	+0.50	-1.00	-1.59	-0.71	-1.40	2.31
Interval	-1.70	-0.06	-1.02	+1.32	-0.73	5.14
MMT'S						
Main	-1.44	-0.52	-0.92	-1.50	-2.20*	0.64
Individuality	-0.48	+1.34	+1.05	-0.75	+0.58	3.36
Interval	+0.72	-0.97	+1.69	-1.00	+0.22	5.25
MGT's						
Main	+0.35	+0.28	+0.12	+0.58	+0.66	0.10
Individuality	-1.30	-0.78	+1.25	-0.77	-0.79	3.80
Interval	+1.19	-1.60	+1.38	-0.60	+0.18	6.21
MMT's, IS						
Main	-1.36	+0.58	+0.09	-1.60	-1.14	3.42
Interval	+1.47	-1.18	+0.18	-1.04	-0.28	4.60
MGT's, IS						
Main	-0.67	-0.35	+0.97	-0.14	-0.09	1.53
Interval	+0.26	-0.74	+1.02	+1.00	+0.77	2.06
MMT's and MGT's, IS						
Main	-1.43	+0.16	+0.75	-1.23	-0.88	3.38
Interval	+1.22	-1.36	+0.85	-0.03	+0.34	3.95

*) $p<.05$, two-tailed
See also note with table 6

completely the accidental finding in the pilot. This caused our significant result for the second of our five formal hypotheses (H2: z=+1.99), a differential effect between observed and hidden PRTs. With z-scores of respectively +0.37 and -2.45, the last, the hidden PRTs, are without doubt to be indicated as the influenced ones. Apart from having this kind of confirmation, we are moreover finding what we are especially looking for, namely a significantly high variance between analyzers (chi-square=9.996; p=.019). This enhanced variance is due to analyzer C producing an extreme result (z=-3.642; p=.0003). Concluding, this analyzer effect makes it plausible that the hidden PRTs are indeed influenced by the analyzers, through the observation the analysis procedure implies, and not by the subjects through some non-observational process such as the transmission of waves or particles.

Why then are the hidden PRTs subject to this analyzer influence? Indeed the observed PRTs show no such effect, but instead a smaller variance than expected by chance (chi-square=0.680; p=.878, see table 6). The explanation for this different behavior can be derived from the hierarchical model we introduced: The model predicts a decreasing susceptibility to the influence of later observers. In this case, the analyzer is the first observer for the hidden PRTs, but for the observed PRTs he is only the fifth. Thereby he has to 'move' the four earlier observers, the subjects, together with the RNG, so that his influence is lesser for the same psi-strength. E.g. if the ratio of the absorbances of psi by observer and observed system respectively is unity, the decrease is by a factor of 16.

The theoretical issue how large this ratio of absorbances is, can be exemplified by the extremes: If the ratio is zero, the observer does not absorb psi and we have again the multiplicative model. On the other hand, if the ratio is infinite, so that the observer absorbs all the psi of the next observer, we have the case where only the first observer determines the outcome of a random event. From the result obtained in our second experiment all we can say about the absorbance ratio is that it is most likely larger than zero.

The other issue of theoretical importance, individuality, is still unsettled; the result for H3 is a non-significant z=-0.57, indicating more positive scoring in the MS condition (z=+0.67; n.s.) than in the IS condition (z=-0.14).

The effect of the feedback time interval in the MGTs of the pilot

experiment ($z=-2.26$; $p=.024$) is not replicated in the MGTs ($z=+0.22$; n.s.) nor in the MMTs ($z=+0.18$; n.s.) of the second experiment.

The data-split for the subjects is conspiciously negative in its results: No result of any subject is significant in table 8 and neither are the between-subject variances.

The findings in the momentanously generated trials of the second experiment deserve an attempt at interpretation, especially in the light of the conclusions on the PRTs. Here we have a feedback condition effect: The MMTs score significantly negative ($z=-2.20$; $p=.028$) and the MGTs slightly positive ($z=+0.66$; n.s.) with a significant difference ($z=-2.02$; $p=.044$). Moreover, this difference is associated with a significant analyzer effect (chi-square=8.44; $p=.038$), making it all the more interesting. A note of caution is appropriate here: The analysis procedure for the MGTs and MMTs was carried out by analyzer D preparing the tables and then sending them to the other three. However, the most extreme scoring analyzer is again C ($z=-2.48$) with D a good second ($z=-2.30$). The sign of these differences between feedback conditions is the opposite from that found in the PRTs, if we compare them straightforwardly: Consider that the MMTs are observed in the experiment together with the observed PRTs, while the MGTs are associated with the hidden PRTs. One would expect in a straightforward view, when considering indeed the subjects as the source of an influence, the differences between the conditions to have the same sign for both types of trials. To demonstrate that the results are at variance with this straightforward view, we test the difference between the effect ($z=+0.37$) in the observed PRTs and its expectation derived from the effect ($z=-2.20$) in the MMTs (see Houtkooper, 1977a). This difference is significant for the expectation derived from the multiplicative model ($z=2.13$) as well as for the hierarchical model with absorbance ratio unity: $z=1.99$; $p=.046$. Thus, in the straightforward view both models are at variance with the data.

With regard to individual differences between analyzers, it may be remarked that C is giving clearly the largest contribution to the analyzer effects, see tables 6 and 7. When the pilot experiment was analyzed, C was present as B was the absentee.

The viewpoint we have developed, and the absence of subject effects strengthens it, is that there is no discernible effect of the subjects in this experiment. This may have an explanation in the inhibiting psychological condition of the experimenters being their own guinea

pigs, with consciousness and creativity only coming into action in the exciting atmosphere of the analysis part of the experiment. Nevertheless, it appears that the analyzers do have an effect. This corroborates much previous work such as Feather and Brier (1968), Taddonio (1976), Bierman and Houtkooper (1980) and Bierman and Weiner (in press). Moreover the analyzers' effect seems all the greater when the subjects haven't been able to make the random events "heavy". This heaviness or decreased susceptibility to psi is, as mentioned above, specifically predicted by the hierarchical model.

The findings in the MMTs and MGTs fit into this picture in the following way: Firstly, the analyzers are the second observers in this case and their influence is therefore diminished according to the hierarchical model by a factor of two. Secondly, the MMTs have only been observed mixed with PRTs, so that the subjects got less information about the MMTs than about the MGTs. This incomplete information in feedback is a phenomenon giving rise to somewhat artifactual effects. E.g. when PRTs are incompletely observed, repeated observations can be regarded as not only influencing the same trials repeatedly, but also as influencing different trials for the first time (Millar, 1978). This results in our case in the MMTs being observed by the analyzers, in effect partly for the first time and partly for the second time, as these had been incompletely observed by the subjects. This argues for a larger influence by the analyzers on the MMTs than on the MGTs, which are all observed for the second time. This fits very well with the results of the second experiment: In the MMTs the main effect is $z=-2.20$ and the analyzer effect is chi-square=7.77, while in the MGTs the main effect is but $z=+0.66$ and the analyzer effect chi-square=3.15. So we have another confirmation of the role that, in this experiment at least, the feedback to the analyzers plays. We conclude that the consistent interpretation of the experiments in this study is only possible through the particular properties of the hierarchical model.

ACKNOWLEDGEMENTS

We are indebted to Dr. Dick Bierman, who brought us together by his efforts to organize the International Summerschool Experimental Parapsychology 1980 at the Research Institute for Psi-phenomena and Physics (RIPP).

We are indebted to Jerry Solfvin and Dr. Brian Millar for valuable and imaginative comments on an earlier draft of this paper.

ABSTRACT

Within the framework of observational theory we introduce a hierarchical model, which regards an already observed random event as a system consisting of the preceding observer and the one level less complex system, both of which act as absorbers of psi. Two experiments were carried out with a repeated retro-PK setup, in which the authors were both subjects and analyzers of the data.

No effect is found with the conditions of individuality, the individual subject observing the same pre-recorded trials (PRTs) 4 times versus each subject observing the same PRTs once. Neither are systematic effects found of the length of the feedback time interval.

A novel condition in this study is to give no feedback about the PRTs. In this condition significant psi-hitting was obtained in the pilot experiment, whereas in the second experiment significant psi-missing is produced. Of the five formal hypotheses in the second experiment, the one significant at the .05 level is the difference between observed and hidden PRTs.

An analyzer split on the second experiment shows significant analyzer effects between the 4 analyzers but no corresponding effects between subjects.

The significant effects in the hidden PRTs of both experiments as well as in the momentaneously generated trials (MMTs) of the second experiment that were mixed with the observed PRTs, together with analyzer effects in the second experiment in the hidden PRTs and in the MMTs, are all found to be in line with the hierarchical model. We conclude that the analyzers, or at least one of them, produced psi, whereas the subjects possibly only inhibited the influence of the analyzers.

REFERENCES

Bierman, D.J. & Houtkooper, J.M. 'Exploratory PK tests with a programmable high speed random number generator', E.J.P., 1975, 1, no. 1, 3-14.

Bierman, D.J. & Houtkooper, J.M. 'The potential observer effect or the mystery of irreproduceability', in Roll, W.G. (ed) 'Research in parapsychology 1979', Metuchen, in press.

Bierman, D.J. 'A test of the potential observer theory using gifted subjects'. in Roll, W.G. (ed.) 'Research in parapsychology 1980', Metuchen, in press.

Bierman, D.J. & Weiner, D.H. 'A preliminary study of the effect of data destruction on the influence of future observers', to be published in the J.o.P..

Feather, S.R. & Brier, R. 'The possible effect of the checker in precognition tests', J.o.P., 1968, 32, 167-175.

Frege G. 'Gedankenfuege', in 'Beitrage zur Philosophie des Deutschen Idealismus', Band III, 1923, 36-51. Repr. in Agnelli, I. (ed) 'Kleine Schriften', George Olms 1976, Huldesheim, 378-394. Transl. 'Compound thoughts', in Geach, P. and Black, M. (eds) 'Logical investigations', Oxford, Blackwell, 1977.

Hartwell, J.W. 'A bound for the observational theories', E.J.P., 1977, 2, no. 1, 19-28.

Houtkooper, J.M. 'A study of repeated retroactive psychokinesis in relation to direct and random PK effects', E.J.P., 1977, 1, no. 4, 1-21 (a).

Houtkooper, J.M. 'A comment on Schmidt's mathematical model of psi', E.J.P., 1977, 2, no. 1, 15-18 (b).

Millar, B. 'The observational theories: A primer', E.J.P., 1978, 2, no. 3, 304-332.

Millar, B. & Hartwell, J.M. 'Dealing with divergence', in Roll, W.G. (ed) 'Research in parapsychology 1978', Metuchen, 1979, 91-93.

Schmidt, H. 'Toward a mathematical theory of psi', J. A.S.P.R., 1975, 69, 301-319.

Schmidt, H. 'A remark on the 'divergence problem'', E.J.P., 1978, 2, no. 2, 163-166 (a).

Schmidt, H. 'Can an effect precede its cause? A model of a noncausal world', Foundations of physics, 1978, 8, no. 5/6, 463-480 (b).

Taddonio, J.L. 'The relationship of experimenter expectancy to performance on ESP tasks', J.o.P., 1976, 40, 107-114.

Walker, E.H. 'The nature of consciousness', Math. Biosciences, 1970, 7, 131-178.

Walker, E.H. 'Foundations of paraphysical and parapsychological phenomena', in Oteri, L. (ed) 'Quantum physics and parapsychology', Parapsychology Foundation, 1975.

Walker, E.H. 'Comparison of some theoretical predictions of Schmidt's mathematical theory and Walker's quantum mechanical theory of psi', J. res. in psi phenomena, 1977, 2, no. 1, 54-70.

Joop M. Houtkooper
Laboratory for Psychophysiology
Eerste Constatijn Huygensstraat 20
1054 BW Amsterdam
The Netherlands

Robin Andrews
20 Tower Road
Hereford HR4 OLF
England

Paul G.J. Ganzevles
Koning Wenzelstraat 1
7415 KK Deventer
The Netherlands

Pieter C. van der Sijde
van Deijsselstraat 24
3319 BW Dordrecht
The Netherlands

CAN EXPERIMENTERS LOCALIZE NONRANDOM TARGET SEQUENCES IN RANDOM NUMBER TABLES?

Trevor A. Harley Carl L. Sargent
University of Cambridge

On theoretical grounds one would expect experimenter psi to contribute some of the variance to observed psi results. If motivation and attitude variables affect psi performance at all, this must be so, since most experimenters are highly motivated 'sheep' in the Schmeidlerian sense.

In this brief preliminary report we detail a replication of an experiment reported by Morris (1968). In that experiment Morris found that he could localize nonrandom target sequences in a table of random numbers which conformed with his 'aim' if he guessed entry points but not if the entry points were prepared by dice-throwing. Allowing for precognitive effects and entry-point derivations which involve some 'psychic shuffle' element this suggests that psi experimenters may indeed be helping the results along by localizing conveniently patterned target sequences in random number tables. It is odd that Morris' experiment has not been the subject of a replication attempt.

In the present experiment we examined only the condition which yielded significance in the Morris study: guessing at entry points. Since most of the derivations of entry points to random tables in experiments in this laboratory use a procedure of the psychic shuffle variety (we are not equipped with an REG at this time) such a procedure has relevance to the possibility of experimenter psi effects in our work.

METHOD

Subjects

There were two subjects in this experiment, the authors of this paper (hereafter TAH and CLS). We have both independently observed significance in psi experiments, CLS more often than TAH simply by virtue of his longer experience in this area. However TAH has generated significant psi-missing from subjects in ganzfeld studies independently of CLS (Sargent, Harley, Lane, and Radcliffe, 1980) and also elicited significant and predicted psi effects in work with a psi-predictive questionnaire independently of CLS (Sargent and Harley, 1980).

Procedure

Each subject independently wrote down 40 entry points to the Rand Corporation (1955) table of random numbers. For the first 20, CLS 'wished for' a high frequency of odd digits (1,3,5,7,9) in the 50 random numbers following the entry point whilst TAH wished for a high frequency of even digits (0,2,4,6,8); for the second set of 20 the aim was reversed. A coin toss by TAH determined the aim order. Subjects checked carefully that they had not written the same entry point down twice.

There is a very recherche artifact here; it is possible that one might subliminally remember particular sequences from the Rand tables and bias one's responses accordingly. For this reason a random transform was added to the second set of 20 guesses for each subject independently. A deck of conventional playing cards with 10, J, Q, K removed (taking A=1) was randomly cut by each subject independently, the denomination of the top card on the cut deck noted, and the value added to the entry points written down. Thus TAH's original entry points read 09998, 10110, ... in the second batch of 20 and after adding a random transform (of +9 in this case) they read 10007, 10119, ..., whilst CLS's original entry points read 07418, 08632, ... and after adding +8 throughout now read 07426, 08640, ...

Subjects now carefully tallied the number of 'target' (odd or even as appropriate) digits in the sequence of 50 following the entry point for all 40 of their entry points. Subjects also checked each other's

data to ensure checking accuracy (as will be seen in the case of the TAH data this is vitally important).

We planned to analyse the mean and variance for each subject and for the combined data and the difference between subjects. We made no predictions because the previous literature simply is not sufficient in terms of amount of data to permit this. Hence all p-values in this paper are two-tailed.

RESULTS

Hit rate

Mean score per 'run' (i.e., number of hits in the 50 digits following each entry point) is 25 by chance. CLS had a mean run score of 23.85; he scored 954 hits, a deviation of -46 which is not quite significant ($t=1.71$, $df=39$, $p=.10$) because the variance of the distribution (18.03) is quite sizeable larger than MCE (12.50), although not significantly so (using the chi-square=(n-1)s**2/sigma**2 and z=sqrt(2*chi-square)-sqrt(2n-2) formulae we obtain $z=1.78$, $p<.10$).

TAH obtained a run-score mean of 26.05 for a total of 1042 hits (deviation +41). This is significantly above MCE ($t=2.01$, $df=39$, $p=.05$).

The difference between the scoring rates is clearly significant (independent means $t=2.59$, $df=78$, $p<.01$).

Variance

The variance for the CLS data was as noted, 18.03 and for TAH it was 10.95. Neither variance differs significantly from MCE (although as noted CLS's variance comes close to doing so) but they are significantly different from each other ($F(39,39)=1.65$; $p<.05$). This seems to be due to the CLS distribution containing several very extreme scores whilst TAH's data contain very few. Thus whilst TAH only has one run score lying outside the range 20-31 inclusive, the CLS data contain run scores of 18, 19, 16, 16, 19, 16, and 17.

Further data breakdown

Table 1 shows the data breakdown by run-group. It may be noted that the variance difference and scoring rate difference between the two subjects are both independently significant only for the second group of 20 runs - i.e., those runs which utilized a transform of the data to eliminate the possibility of subliminal recall influencing the scoring rate. Thus this finding strengthens one's faith in the nonartifactual nature of the findings.

TABLE 1
Breakdown of data

	Entry points 1-20		Entry points 21-40	
	Mean	Variance	Mean	Variance
TAH	26.35	13.47	25.75	8.75
CLS	24.50	17.42	23.20	18.69
Difference means	1.85		2.55	
t(diff), 38 df	1.49		2.18*	
F(19,19)		1.29		2.14*

*) $p<.05$, two-tailed

DISCUSSION

The findings of this study were not predicted in advance of data collection and hence must therefore be treated with the appropriate reservations. However, it was observed that the two psi experimenters showed significantly different scoring in the task both in terms of hit rate ($p<.05$) and variance ($p<.05$). Moreover these effects were found to be independently significant only in that portion of the data which could not contain an artifact possibility (although the effects here were not significantly different from those found in the rest of the data and a temporal incline effect is confounded with the difference in conditions as regards an artifact possibility).

The fact that the significance takes the form of a differential across experimenters is probably not surprising. We have worked closely as a pair in the context of ganzfeld free-response GESP research for over a year and a half and close co-operation in such settings tends to 'blur the edges' of individuality; experimenter teams (in our experience at least) tend to become closer to Gestalts after a while and one of our problems in the ganzfeld work is that when we work as sender and experimenter, with a third person as subject, there is not infrequent psi contact between us which gives the experimenter problems when he comes to help the subject with the judging.

This report is a preliminary one. It's major aim is to push researchers into examining further the possibility of experimenter psi-mediated localization of desired sequences in random number tables. If one such empirical enquiry is motivated in part by this report, our job is done.

ABSTRACT

The two authors of this paper attempted to guess entry points to the Rand Corporation tables which would yield 'aimed-for' targets (odd or even digits, aim being counterbalanced within and across subjects). In the case of half the data an arbitrary addition to the subject's written entry points was made to eliminate the possible arifact of subliminal recall of the tables influencing calling.

CLS obtained marginal ($p=.10$) psi-missing and high variance ($p<.10$). TAH obtained significant ($p=.05$) psi-hitting and low variance. The two subjects showed significantly different performance in terms of both scoring rate ($p<.01$) and variance ($p<.05$). Both these effects were only significant independently for the half of the data in which the arbitrary addition noted had been made, which was thus artifact-free.

The results suggest that further exploration of the possibility of experimenter psi-mediated localization of 'conveniently patterned' target sequences should be made.

REFERENCES

Morris, R.L. 'Obtaining non-random entry points: A complex psi task', in Rhine, J.B. and Brier, R. (eds) 'Parapsychology Today', New York, Citadel, 1968, 75-86.

Rand Corporation 'One million random digits', Glencoe, Free Press, 1955.

Sargent, C.L. & Harley, T.A. 'A simple psi-predictive questionnaire for mass testing', Paper presented at the Fourth International Society for Psychical Research Conference, Brighton, 1980.

Sargent, C.L., Harley, T.A., Lane, J. & Radcliffe, K. 'Ganzfeld psi-optimization in relation to session duration', 23rd PA Convention, 1980.

Trevor A. Harley
Carl L. Sargent
Psychological Laboratory
University of Cambridge
Cambridge CB2 3EB
United Kingdom

EMOTIONS AS TARGETS IN ESP STUDIES

Sybo A. Schouten
University of Utrecht

Although most researchers in parapsychology are likely to agree with the idea that emotions play an important role in ESP the evidence for the validity of that notion is rather scanty and largely anecdotical. In fact so little ESP research applying this concept has been carried out that the word emotion does not even appear in the subject index of the 'Handbook of Parapsychology' (Wolman, 1977).

The assumption of a relationship between emotions and ESP probably originates from the study of spontaneous cases where in all analyses it is found that in the majority of the cases emotionally rather dramatic events have taken place. The idea is also rather popular amongst paragnosts or mediums of which many believe that only events which are emotionally important to a person are accessible by ESP. Although not based on systematical studies such beliefs, when held by persons who are supposed to have frequent first-hand experience with the paranormal, will undoubtedly enhance the popularity of the idea. Still another source for linking emotions with the paranormal is the concept of the unconscious. ESP is supposed to be an unconscious process. Viewed psychodynamically, unconscious processes and contents are considered to be strongly emotionally charged. Hence the implicit assumption can easily be formed that the ESP content which becomes conscious must also have such an emotional loading.

Assuming that, in the absence of supportive research data ,these considerations are strong enough to suggest a relationship between

emotions and ESP it is not easy to specify the nature of such a relationship. This is not only because the concept of emotion is rather confusing and used in a variety of ways but also because it might play a variety of roles in cases of ESP. In spontaneous cases, for instance, events like dying, being involved in an accident, etc., undoubtedly have a strong emotional impact on the target person. In many instances it might be that this emotion triggers the ESP experience in the percipient. But it could also be that the nature of the relationship between the persons is the crucial factor and that this relationship must be of an emotional character. Another possibility is that the events which have a potential emotional effect on the percipient are the ones which are most likely to become the subject of an ESP experience. Hence for spontaneous cases alone one can already distinguish several possible ways in which the concept of emotion can be associated with the ESP experience.

Not only does the possible relationship of emotion with ESP vary, but the very meaning and content of the word emotion varies also. When speaking about an emotional relationship the word emotion means something different than when using the word when speaking about an emotional event. Furthermore the fact that one is involved in a process of dying or in an accident will yield undoubtedly an entirely different type of experience than a situation in which all environmental stimuli are excluded, such as in the case of a percipient. Hence one might ask what exactly do we mean when using the word emotion and to what situations do we apply this concept. As to the latter I will restrict the discussion especially as far as models on emotions are concerned mainly to those models which are of relevance to the situation of someone who either is percipient in a spontaneous case or who acts as a subject in a free-response study. Both cases are characterized by an absence of external perceptual stimuli which normally do play a role in emotions.

In experimental psychology a considerable amount of research has been devoted to this topic. One can, according to Bindra (1970), distinguish between four main lines of experimental work, each attacking a different aspect of this complex concept. These are: (1) the study of emotional experience; (2) of emotional arousal: the changes in internal bodily processes; (3) of emotional action: the overt environmental directed response patterns; and (4) of emotional stimulation: the particular stimulus features of the environment that usually provoke one or more of the above three. Thus, in general, emotion is viewed as a concept that has neurophysiological,

motor-expressive and phenomenological aspects (Izard, 1972).

Studies in experimental psychology have mainly concentrated on the neurophysiological and motor-expressive aspects of emotions, probably because the first two rivalling theories about the nature of emotions disputed the role of visceral-somatic reactions. William James held the opinion that these reactions follow directly the perception of the exciting stimuli and that our experience of these bodily reactions constitute the emotional experience. Cannon leveled several criticisms at this visceral response hypothesis and defended a central neural model in which the visceral reactions and subjective feelings are both a consequence of central processing of emotional stimuli. It is now generally accepted that autonomic reactions are too unspecific to be considered as the source of different feelings, but that they somehow play a role (Schachter & Singer, 1962). This was one of the arguments used to construct a cognition-arousal model. That model assumes that arousal in itself does not define a clear state of emotion but that it forces the subject to scan his environment to locate a cause for the arousal and when such a cause is identified this cognitive aspect will lead to the particular emotional experience. "It is the cognition which determines whether the state of physiological arousal will be labeled as anger, joy, fear, or whatever" (ibid, p.380). Further on they state that emotion is "a joint function of a state of physiological arousal and an appropriate cognition" (ibid, p. 382). This 'two-component' model could explain why we are able to interpret emotional feelings in others. We do this by combining the knowledge of the visible effects of our own visceral reactions in emotional states, like sweating or turning red, and interpreting cognitively the situation the other is involved in.

Whatever the merits of these models there seems to be a general acceptance of the association between emotional experience, certain stimuli, arousal, and cognitive appraisal. This is reflected in the nature of the experimental research in this area. One would expect in studies on emotions with human subjects the emotional experience to be the dependent or independent variable, or at least to function as an indicator for the occurrence of emotions. However, such is hardly ever the case (see for instance Levi, 1975). In most studies external stimuli are presented and it is assumed that as a reaction to these stimuli subjects will experience emotions. Sometimes autonomic reactions are registered and applied as indicators of an emotional experience. That such operationalizations are generally applied must be based on the assumptions that not only arousal and certain stimuli

are associated with emotional experiences but also that either such stimuli or arousal would be a sufficient condition for eliciting emotional experiences. Especially the latter assumption seems in many instances doubtful.

One important consequence of viewing emotional experience as dependent on certain stimuli, arousal, and cognitive appraisal is that it becomes unavoidable to consider these experiences as the outcome of a secondary or categorizing process. Thus only after experiencing external stimuli, changes in arousal, and after interpreting cognitively our situation are we able to label the total experience and to say, for instance, that we feel angry. This view implies that in fact emotional experiences can not exist as individual separate experiences which can be immediately identified based on the experience itself such as when we identify pain. According to this model we experience arousal and only after interpreting our situation are we able to label the emotion in certain 'emotional' terms. Consequently, without the possibility of cognitive appraisal and/or without experiencing arousal, we would not be able to have an emotional experience, and hence it's secondary character.

Although we can assume that sometimes emotional experiences are associated with changes in arousal, one could argue that the model which treats the emotional experience as fully dependent on arousal and cognitive processes is to a certain extent a consequence of the way emotions are operationalized in experimental research. Emotional experiences are strictly individual and not objectively measurable. Hence it becomes necessary to apply some operationalization which can be measured objectively. However, when a certain operationalization is applied the phenomenon and its operationalization tend to become mixed up and to be considered as more or less the same thing. It is not surprising, since changes in arousal can be measured objectively, that emotional experiences and arousal are often treated as being the same phenomenon.

Since so few studies have been aimed at finding the circumstances under which people do experience emotions and considering the rather far-reaching implications I feel it justified to question the assumption on which the views discussed above are based that emotional experiences are basically a matter of cognitively interpreting an experience of a different state of arousal. It might be that people also experience emotional feelings under quite different circumstances or that situations in which we should expect emotional reactions do

not result in such experiences at all. At least it is not difficult to find anecdotical examples for such exceptions.

For example, people do experience emotions in the absence of arousal when feeling happy or when being in a prolonged state of depression. On the other hand, people can show strong arousal without experiencing emotional feelings such as in vigilance tasks. People can feel strongly emotional even in the absence of both external stimuli and cognitive appraisal such as when waking up from a nightmare and feeling very emotional without knowing the reason. The picture is further complicated by the well-known phenomena of (classical or operant) conditioning and adaptation. A person can become conditioned to respond emotionally to specific (internal or external) stimuli which normally do not elicit such reaction while similarly by adaptation someone can learn not to respond emotionally to stimuli which normally evoke emotions. One might even wonder whether emotional experiences are really that dependent on cognitive appraisal considering the behavior of young children and animals.

This rather superficial discussion of these aspects of emotional experiences is presented here since they have a direct impact on the issue of the relationship between ESP and emotions. If emotions are a secondary process, a matter of experiencing arousal and cognitive appraisal of one's situation, then the possible relationship between emotions and ESP can hardly be of importance. If external stimuli are considered necessary for eliciting emotional experiences then verbal reports of emotional experiences under circumstances where such environmental stimuli are lacking - such as in spontaneous ESP experiences or in the case of a subject in a ganzfeld study - become evidently meaningless.

However, when it is recognized that emotional experiences can also be initiated by intra-individual processes, a position taken, for example, by Izard (1972), then one could perceive a model which puts the emotional experience in a central position while giving the phenomena of arousal and cognitive appraisal an associative and secondary status. Such a model would treat the emotional experience as a primary process which allows us to consider a more meaningful relationship between ESP and emotions.

Without trying to do so I assume for argument's sake that such a model could be developed. At least as far as I can judge it would not contradict the few scanty data on emotional experiences we have. One

of the first questions which arises then is to what extent people are able to discriminate between emotional experiences based only on the experiences and which variables are relevant for the discriminatory process. From a parapsychological point of view these questions are of primary importance. Emotions can only be of importance in the study of ESP when people are able to discriminate between various emotional experiences purely on basis of the experience itself. If that is not the case, as in Schachter's previously mentioned two-component theory, then other information has to be acquired first before someone can label the associated experience in emotional terms. If, however, emotions can be experienced as a distinct state for which knowledge of the environmental situation is not required for its interpretation then one can imagine that such states are 'sensed' and correctly interpreted by for instance a percipient or a subject in an ESP experiment.

In ESP experiments the majority of targets are visual, both in forced-choice and in free-response studies. However, it is a somewhat peculiar combination considering ESP in terms of an information-processing model and at the same time assuming implicitly that elements belonging to a clearly different information-processing system would be most optimal to be transferred by ESP. It might be comparable to a situation in which mainly auditory stimuli would be applied to study the olfactory system. Thus we may ask, with respect to the ESP process, whether emotional experiences seem worthwhile candidates to consider as targets.

From spontaneous cases one is tempted to conclude that the percipients do not 'see' what happened to the target person but that they somehow translate a few elementary elements of information about the target person in a more complicated 'picture' of own fabrication (Schouten, 1978). In a lot of these cases the pieces of information (death, drowning, accident, etc.) are such that they could be deducted from the associated emotional experience. For instance an experience of despair and suffocating might indicate drowning. Consequently one could device a model of ESP in which not quasi sensory information but emotional experiences or feelings are the primary sources of transfer. But again such a model is only feasible when it can be shown that emotional experiences can be considered as a primary process and that people can discriminate between these experiences without knowledge of situational circumstances.

This study is devoted to the problem of whether people are able to

discriminate between emotional experiences only on basis of the experience itself. An essential aspect of the study is that these experiences have to be confined to 'neutral' environmental situations in order to assure that external stimuli will not play a role in making the discrimination. They should be studied purely as experiences on their own.

A literature survey revealed that only few studies could be found aimed at studying emotional experiences directly. These studies, for instance Izard (1972) and Davitz (1970), have been mainly of the type of factor-analyzing descriptions of subjects given in non-emotional states of different emotions. Such studies result in various 'list' of emotions and associated words. One conclusion becomes quickly apparent. Namely that not only environmental appraisal but also verbal components play an important role in 'labeling' emotions.

Hence the problem whether and how people discriminate between emotional experiences in the absence of external stimuli becomes undistinguishable from the problem of how people communicate about these experiences. Even if people can experience different emotions in this condition they should be able to communicate about it in such a way that amongst subjects the same words refer to similar experiences. It is of course impossible to check this in a direct way but we can get an idea of the degree of similarity by asking people to describe emotional experiences in non-situational terms. If we would find that for a specific emotional experience all subjects apply a unique set of adjectives then one might assume that the experiences behind these descriptions are rather similar. In fact I will base the interpretation of the results of this study on the premise that similar descriptions between subjects of the same emotional experience is sufficient proof for the existence of such an experience. If however people describe different experiences in similar terms or in different terms but in an inconsistent way then at least from a practical and experimental point of view it becomes meaningless to study these experiences or to make them a vital component of an ESP model. This, of course, is provided that no other methods could be developed to discriminate between such experiences.

Hence this study was directed at the question whether people apply the same or different descriptions for various emotional experiences and how they compare these emotions with other emotions when asked to experience and rate these emotions in the same neutral environmental context. The situation in which these emotions are experienced is thus

in fact comparable to the situation of an ESP-subject in a free-response study or to the situation of a percipient in a spontaneous case who apparently without environmental stimulation suddenly experience a strong emotion. For practical purposes it was also hoped for that as a result of this study a method could be constructed with which types of emotional experiences could be assessed based on the existence of a unique set of adjectives for each experience.

METHOD

An important problem in studying emotional experiences when environmental stimulation is to be held neutral is to elicit such experiences. This problem was attacked by requesting subjects for each emotion to recollect and 'relive' the last time they had experienced that emotion strongly. In other words they had to recreate the whole episode leading to the particular emotion in their minds and to experience again as much as possible all feelings they had at the time. Then when they felt they were involved again they described on a form in their own words how they would typify the experience. Then they circled from a list of 47 adjectives not related to situational aspects those adjectives which they considered appropriate for that experience. Next they were asked to indicate which of nine other emotional experiences they found most similar to that experience and to indicate those they had experienced more strongly in their life. The last question asked whether they considered that experience as positive or negative. To avoid an influence from order effects the questionnaire was constructed in such a way that each subject was presented with a different order in which to experience and to rate the emotions.

Since emotional experiences and the circumstances under which they occur are highly private matters anonimity was assured by distributing the questionnaires via third parties and by omitting all questions about personal data like age, gender, etc.. Consequently no data are available on the properties of the sample. However, subjects were invited the way normally people would be attracted to take part in ESP studies so to a certain extent the data can be considered as representative for this group. It was expected that subjects would devote time on the experiment when being alone and when they felt like it. No return date was given. Subjects were not offered any rewards. It was stressed that taking part in the experiment would require

several hours and that in view of the importance of obtaining reliable data it was considered preferable that they not return the questionnaire than to return it with parts of it filled in quickly or superficially (because of boredom, etc.).

From the few studies available ten emotions were selected which can be considered to belong to the group of most basic emotions. These emotions were joy (1), depression (2), repulsion (3), shame (4), in love (5), grief (6), jealousy (7), anger (8), freedom (9) and fear (10). The questions about these emotions were formulated in a direct way, for example: "Did you experience the feeling of being angry as positive or negative?" When comparing experiences the other nine were all listed in random order for marking purposes.

RESULTS

Fourty-eight subjects completed and returned the questionnaire. The descriptions of the emotions given in their own words were generally short and varied extensively in content. Although occasionally in those descriptions a dominance of some adjectives with some specific emotion was noticed no consistent pattern could be found typifying each emotional experience. Since these adjectives also appeared in the list of 47 adjectives presented on each form the higher frequencies observed might have been influenced by the repeated confrontation of the subjects with these list of adjectives.

For each of the emotions the frequency distribution for the 47 adjectives over all subjects was fixed. Then for each emotion the distribution of the adjectives was compared with the distribution of adjectives for each other emotion by applying two different tests. In the first place for each of these comparisons the cumulative distributions were formed and tested for similarity by applying the Kolmogorov-Smirnov two-sample test (Siegel, 1956). It was decided that the distributions would be considered not to differ when the probability level of the computed chi-square value for the maximum difference between the distributions would be higher than $p=0.10$. Secondly because an absence of significant difference does not automatically imply a relationship a product-moment correlation was computed for each pair of distributions. Considering the type of data this procedure is not entirely correct, but it does give us an indication as regards sign and strength of relationship. Table 1

presents the results for those pairs of distributions of adjectives who according to the Kolmogorov-Smirnov test do not differ.

TABLE 1
Comparison between pairs of emotions for which the distributions of verbal descriptions do not differ.

Pair	Dmax	r	Pair	Dmax	r
1-9	.04	.90	2-6	.09	.85
1-5	.10	.90	2-10	.10	.39
5-9	.11	.78	6-10	.11	.28
4-7	.12	.59	6-7	.13	.47
2-7	.13	.47	4-10	.13	.60
2-4	.14	.42	7-10	.16	.51
4-6	.17	.29	7-8	.18	.65
3-8	.21	.26	5-8	.17	-.18
3-5	.20	-.55			

When considering the full matrix of product-moment correlations (not presented here) it becomes immediately apparent that all negative correlations (except one zero-correlation of -.004) are between all combinations of elements of the group 1,5,9 with elements of the group 2,3,4,6,7,8,10, while within these two groups all correlations for pairs are zero or positive.

From table 1 it can be concluded that except for the combinations 5-8 and 3-5 the frequency distributions for the different pairs of emotions listed are positively correlated. Since we know that these distributions do not differ significantly and that they are positively correlated we can try to form clusters of emotions of which all elements are associated with all other elements. This shows to succeed quite easily. The first cluster is formed by the emotions 1-5-9 of which the associated distributions of adjectives correlate strongly amongst each other but not with the distribution of adjectives of any

of the other elements. The second cluster is formed by the elements 2-4-6-7-10 of which all possible combinations can be found in table 1. The remaining two combinations which do not fit into a full cluster are 7-8 and 3-8. Thus based on the adjectives associated with each emotional experience we can not distinguish between the emotions joy, love and freedom. The same applies to the emotions depression, shame, grief, jealousy and fear.

The subjects were also requested to indicate those emotions which in their opinion resembled most the experience under consideration. They were free to mark none or any number. Contrary to the previous analysis this question was directed at comparing the experiences themselves. Discrepancies between pairs are to be expected, for instance with emotion 1 emotion 5 was marked 29 times as resembling emotion 1 but emotion 1 was marked 41 times as resembling emotion 5. I do not consider such discrepancies as due to errors. The nature of the experiment (the many hours it takes, etc.) makes it unavoidable that the context in which such judgments have to be made will differ for the various emotional experiences. Hence they were presented in a randomized order.

For each pair of emotions the number of times the other was marked as being a similar experience was added for both elements. The resulting matrix is presented in table 2.

On this matrix a simple cluster analysis was carried out by applying the following procedure. The pair with the highest total frequency (thus with the greatest similarity) was taken as starting point. Then another element is selected in such a way that the average value of the frequencies of all combinations is maximal. In other words, the elements are on the average most similar. The standard deviation for these frequencies is computed and the average frequency of the elements of the cluster is compared to the highest frequency of an element of the cluster and any other element. If the observed difference between average cluster frequency and highest frequency of a combination with a non-cluster element does not differ considering the size of the standard deviation then the procedure is repeated with another element. In the case the highest frequency of a combination with a non-cluster element shows to be significantly lower than the average cluster frequency the cluster is considered definite and a new search is started based on the highest frequency between two new non-cluster elements.

TABLE 2
Number of times elements of pairs of emotions were marked as being similar experiences.

Emotion	1	2	3	4	5	6	7	8	9	10
1										
2	1									
3	0	13								
4	0	19	30							
5	70	0	0	1						
6	3	76	9	16	5					
7	0	10	11	12	3	28				
8	0	7	46	10	1	20	24			
9	84	0	0	0	31	0	0	1		
10	0	26	12	24	0	25	16	17	1	

Starting with 9-1 (f=84) the highest average frequency value is found when adding element 5. The average value becomes 1/3 * (84+70+31) = 61.6, with an associated standard deviation of 27.5. The highest frequency between any cluster element with any of the remaining elements is found to be 1, which can be considered to be significantly different from the value 61.6. Hence the elements 1-5-9 (joy, love, freedom) can be considered to form a cluster.

The procedure is repeated starting with the combination 2-6 (f=76). Table 3 presents the highest average frequency, standard deviation and highest frequency of an element of the cluster with another element for each step of the procedure. As can be seen the cluster found above based on analysing the frequencies of adjectives (2-4-6-7-10) emerges in the sixth step. However in this case the procedure has to be continued and only ends after all remaining elements have become part of the cluster. It is also apparent that the average cluster values decrease gradually and that nowhere a clear cut-off point appears which might present a reason to consider this cluster as an

association of two different clusters.

The analysis of the data acquired from the answers on the question which emotions the subject had experienced stronger than the emotion under consideration is based on the well known method of paired-comparison. The data are treated as if the subject compared each emotional experience with each other emotional experience judging which of the two he considerd a stronger experience. In reality the subjects were only asked to mark those emotions they had experienced as stronger.

TABLE 3

Average frequency, standard deviation and highest frequency with a non-cluster element for all steps in the cluster analysis.

Combination	Average frequency	S.D.	Highest frequency to non-cluster element
9-1	84		70
9-1-5	61.6	27.5	1
6-2	76		28
6-2-10	42.3	29.2	28
6-2-10-4	31.0	22.4	28
6-2-10-4-7	25.2	18.9	20
6-2-10-4-7-8	22.0	16.3	46
6-2-10-4-7-8-3	20.4	15.2	1

As a consequence they were not forced to give markings in the case they considered both emotional experiences as equally strong. For the evaluation it was assumed that when the data for all subjects were combined the no-answers for each pair indicated no difference in strength. For instance 28 subjects rated emotion 1 as stronger than emotion 7 while 3 subjects rated emotion 7 stronger than emotion 1. Since no inconsistencies within subjects were observed e.g. the same subject rating emotion 1 stronger than emotion 7 and emotion 7 stronger than emotion 1 it was assumed that the remaining 48 - (28+3) = 17 subjects could not discriminate between the strength of emotions

1 and 7. Hence these subjects were equally distributed over both cells, thus 36.5 subjects being considered as judging emotion 1 stronger than 7, and 11.5 subjects being considered as judging emotion 7 stronger than 1.

The scale construction yielded three clusters of emotional experiences with approximately equal scale values. As most strongly experienced were rated the emotions 5 (love; s=.60), 1 (joy; s=.47), 10 (fear; s=.40) and 6 (grief, s=.38). The next cluster consists of the emotions 8 (anger; s=.08), 9 (freedom; s=.01) and 2 (depressed; s=-.05). The emotions less strong experienced appears to be 4 (shame; s=-.54), 7 (jealousy; s=-.72) and 3 (repulsion; s=-.76).

The last question was related to whether the subject experienced the emotion as positive or negative. In most cases over 90 % of the subjects responded in the expected direction (for instance love: positive, fear: negative). Although in the remaining cases there is a clear majority for rating the experiences as negative the difference for shame (positive-5, negative-38), grief (positive-12, negative-29) and anger (positive-15, negative-29) is not so striking. The more or less therapeutic value of these experiences might have influenced these judgments.

DISCUSSION

Although of the ten emotions investigated only three are considered positive experiences it is reassuring that in the group of four emotions that our subjects experienced most strongly, two are of a positive nature. In fact these two (love and joy) received the highest scale values and thus were rated as the emotions experienced most strongly.

The difference between positive and negative emotions plays an important role in both types of cluster analyses carried out. Based on the analysis of verbal descriptions as well as on comparing emotional experiences directly it was found that the positively rated emotions cluster strongly and that they are not related to any of the others. The same applies to the negative experiences. Based on the frequency distribution of the adjectives it is not possible to discriminate between depression, grief, fear, shame and jealousy. Based on the direct comparison of the similarity of experience, all negative

emotions form one cluster. It follows that, based on the descriptions given by the subjects, one can reliably distinguish between positive and negative experiences, but not between experiences within each of these groups. One may therefore conclude that people do have similar experiences but only of a positive or negative nature, and secondly that the further labeling of experiences, that is the more refined discriminations within the positive or negative group, must depend on factors different from the experiences themselves.

As regards the main purpose of this study these conclusions imply that from a practical point of view the value of applying emotional experiences as targets in an ESP study or as elements of an ESP model must be considered to be doubtful. We are only allowed to speak of positive or negative emotional experiences. Unless a better method is found (if such is possible at all) which enables us to discriminate between various emotional experiences I feel that further research along these lines is not advisable.

It should be stressed again that the data for this study were gathered in an environmental situation characterized by an absence of external stimuli who could provoke emotional feelings and by an absence of other human beings to express emotional feelings to. Hence one could raise the objection that under these circumstances the subjects possibly did not experience emotional feelings at all. At least according to a cognition-arousal model the lack of arousal stimuli and the impossibility in such an 'empty' environment of attributing cognitively possible feelings of arousal to a specific situation should both prevent the occurrence of emotional experiences. Even if this is assumed to be true it would not necessarily invalidate the main conclusion of this study. If asking subjects to remember highly emotional episodes in their life and to recreate those episodes in their minds does not lead to any emotional feelings then it is also not to be expected that emotional feelings can be experienced in a similar experimental situation when carrying out an ESP study.

As regards the implications of the findings of this study for a general model of emotions I feel that the issue whether the subjects in this experiment (or in other experiments along these lines) did or did not experience emotional feelings is not of much relevance. Emotional experiences are strictly individual matters and can only be studied by eliciting communications about them. The subjects in this study behaved as if they had experienced the different emotions, either at the time of filling in the questionnaire or in the past, and

answered the questions about them as if they could relate these questions to specific experiences. It seems to me a purely verbal matter whether one is or is not allowed to say whether they really did experience something.

At least we can conclude that based on the descriptions given by the subjects it is possible to discriminate between positive and negative experiences. As stated in the introduction this implies that considering the agreement between all subjects as regards these experiences we may assume that such experiences exist and that one might conclude that people are able to discriminate between positive and negative experiences based on the experience itself. This conclusion contradicts a cognition-arousal model which assumes emotional experiences to be the result of a cognitive process elicited after having noticed a state of arousal in the organism. It should be kept in mind that arousal in itself is a neutral concept and that being in a state of arousal does not automatically imply a positive or negative experience. In many situations we can be highly alert (e.g. traffic) without associating this state with positive or negative feelings.

But although it is apparently possible to discriminate between positive and negative experiences directly the fact that the verbal descriptions within each group do not differ forces us also to conclude that in discriminating between experiences within each group situational and/or verbal components do play a major role. This conclusion can be used to support a cognition-arousal model. However, since in this study it was found that people can experience positive or negative feelings without cognitive appraisal the interpretation of this finding can also be a different one.

In a cognition-arousal model both arousal and the labeling process are conditional on being able to experience an emotion. Based on this study I would like to propose that experiencing a positive or negative state is conditional for having an emotional experience. If that assumption is correct it implies that both arousal and cognitive appraisal must be considered as phenomena resulting from the emotional experience. As a consequence cognitive analysis of the situation in which the emotion was experienced can be considered as an effort by the subject to understand what is happening and the resulting labeling of the experience can be considered as having mainly a communicative purpose. The labeling makes the whole experience understandable both to the subject as well as to the others. As regards the latter aspect

it follows that in this view the verbal labeling does not principally differ from expressive behavior in animals or non-verbal expressions like motor-facial activities in humans.

It is assumed that emotional experiences are based on positive or negative experiences which are associated with mental states and which are not merely a reflection of the degree of physical well-being. Arguments for this assumption can be found in the common experience that people do show and report positive emotions when being in physically bad states and vice versa. Moreover people often express that they feel not well, a negative experience, because for instance they suspect that they have catched a flue. This shows that it is possible to discriminate between negative feelings following physical causes and negative feelings considered as emotional based on the experiences alone.

Hence we can construct a model for emotions which deviates from a cognition-arousal model mainly in that it assumes that people do have primary emotional experiences which can be either of a positive or negative nature. They are experienced as emotional feelings and as different from feelings resulting from the state of physical well-being of the organism. These emotional feelings are elicited in a variety of situations both by internal or by external stimuli, and very probably conditioning plays an important role in associating positive and negative emotional feelings with certain stimuli. Thus it is to be expected that to the subject it is often not possible to locate the immediate cause for feeling positive or negative. In that case one might expect a cognitive process to start in which the subject looks for a probable explanation for his feelings. The emotional experience will be followed by arousal in situations which are by nature associated with taking action by the organism (for instance fear - desire to run; anger - desire to attack; but no arousal in the case of depression which is asociated with a tendency to withdrawal). The labeling of the situation enables the subject to communicate his interpretation of the situation and his feelings to others.

This model explains many aspects of emotion. People can feel strongly emotional even when misinterpreting the situation, a phenomenon known as misattribution, because the emotional experience depends on stimuli other than the cognitive appraisal of the situation. Hence one can imagine someone 'exploding' when something innocent is said, a phenomenon dificult to explain in a

cognition-arousal model. The model also allows us to interpret emotional feelings in others correctly, because we are able to share positive and negative feelings since we experience them ourselves and because we are also able to interpret the situation the other is involved in as similar to situations we experienced ourselves. That people do discriminate between different positive and negative feelings although according to their verbal descriptions the experiences themselves apparently do not differ becomes understandable when we accept that the different names for either positive or negative feelings have mainly the function to tell us something about the circumstances in which the feelings were experienced. In animals one could view the function of expressing emotional behavior as communicating intended behavior. With human beings the function of expressing emotional feelings is mosttimes a similar one, its main purpose is to make us understandable to others. Verbal labeling is superior to other ways of expressing emotions in the sense that it communicates in one word a lot of information both about the type of feeling (positive or negative) as well as about the circumstances. If the hundreds of words existing which are related to emotional states would indeed reflect different individual experiences then the use and interpretation of these words must resemble a sort of random process. However, if we interpret the use of these words as revealing something about a situational context then it is not surprising that people do use all these different words in a meaningful way and without creating confusion.

If the views on emotions presented above are correct one can see some practical implications for the study of emotions and arousal. One is that since arousal indicates a neutral state of activation which is not associated with positive or negative feelings subjects should be able to discriminate between pure arousal states and states in which they have emotions. This means that it must be possible to find a method to discriminate between such emotional and arousal experiences. It should also be possible to predict which emotional experiences will be associated with arousal and which ones not. Another implication would be that in order to study emotional experiences it is sufficient to measure whether people experience significantly stronger positive or negative feelings during the experimental session than in general. Here again it does not seem too difficult to devise a reliable test to make such measurements. Perhaps applying such tests when studying emotions experimentally would reduce a bit of the confusion which presently is associated with the concept.

ABSTRACT

Although many parapsychologists suppose that a relationship exists between emotions and ESP there is hardly any experimental evidence to support this hypothesis mainly because experiments using emotions as either targets or variables are rare. This is undoubtedly partly due to the confusion which is associated with the concept of emotion and the resulting lack of appropriate measurement methods.

Emotions are primarily subjective experiences which, under normal everyday circumstances, are associated with psychophysiological, motor-expressive and environmental variables. Since it is difficult to study subjective experiences most studies on emotions focus on the role of the associated variables and consequently most current models on emotions consider the subjective experience to be dependent on the occurrence of such associated phenomena.

Particularly the assumed dependence of emotions on environmental cues presents a problem for parapsychological work. If emotions are to serve as targets in an ESP experiment subjects should be able to discriminate between different emotions strictly on the basis of the subjective experience since in ESP studies external sensory stimuli are excluded. Hence the aim of the present study was to find out whether subjects are able to distinguish between different emotions under such conditions, and whether they are able to communicate their experiences in a reliable and consistent way.

Based on an investigation with 48 subjects it was concluded that subjects are able to discriminate between positive and negative emotions, but not between the different emotions within these two categories. It is therefore suggested that the range of emotions, beyond the positive-negative breakdown, may not be useful as targets in an ESP study.

A model is presented which assumes that people experience only positive and negative emotional feelings, and that the further discrimination between different emotions depends on environmental variables and mainly serve a communication purpose.

REFERENCES

Bindra, D. 'Emotion and behavior theory: Current research in historical perspective', in Black, P. (ed) 'Physiological correlates of emotion', New York, Academic Press, 1970.

Davitz, J.R. 'A dictionary and grammar of emotion', in Arnold, M. (ed) 'Feelings and emotions: The Loyola symposium', New York, Academic Press, 1970.

Izard, C.E. 'Anxiety: A variable combination of interacting fundamental emotions', in Spielberger, C.D. (ed) 'Anxiety: Current trends in theory and research', New York, Academic Press, 1972.

Levi, L. 'Emotions: Their parameters and measurement', New York, Raven Press, 1975.

Schachter, S. & Singer, J.E. 'Cognitive, social and physiological determinants of emotional state', Psychology Review, 1962, 69, 379-400.

Schouten, S.A. 'Analysis of spontaneous cases as reported in 'Phantasms of the Living'', E.J.P., 1979, 2, 4, 408-456.

Siegel, S. Non-parametric Statistics. McGraw-Hill, 1956.

Wolman, B.B. 'Handbook of parapsychology', New York, Van Nostrand Reinhold, 1977.

EXPLORATORY OBSERVATIONS WITH PK METAL BENDERS FROM SOUTHERN CALIFORNIA

Mark G. Shafer
McDonnell Laboratory for Psychical Research

BACKGROUND

This paper summarizes observations from two years' exploratory research on PK metal bending in a small group setting, with selected persons from Southern California. A total of 13 people participated in the research, from 1976 to 1978. Participation was voluntary and by the invitation of the experimenter, the author, only.

Participants ranged in age from 12 through 40, with 5 being under 20. All claimed macro-PK ability, and were invited to participate in the research either after making this claim at one of the author's public lectures, or after being referred to the author by an acquaintance familiar with his research.

All participants claimed to be able to bend stainless steel spoons and/or other cutlery psychokinetically by holding the item in one hand and rubbing the neck continuously between bowl and stem until through PK the metal softened there. At this point a bend or twist would be attempted using the other hand for physical manipulation. Occasionally such attempts appeared very effortless, involving no more than a quick, light flip of the wrist. Participants 'knew' when a spoon was softened either through periodic testing with the other hand 'to see if it was ready to go', or through feeling it become slippery beneath the grasp, 'as if suddenly coated with mercury'. The psychokinetically

softened condition was reported to last only a fraction of a second.

Participants all claimed to have noted the capacity for PK after hearing of, or seeing a demonstration of apparent genuine metal bending either in person or on television. Each later tried similar PK feats and at some point met with what they took to be success. Some continued to practice after their apparent success and others desisted from further attempts at this time. At some point during the research all but one of the participants mentioned having a relative or some significant person in their past who professed a belief in psi and who influenced their thinking about it, and all but two professed the belief that other members of their family or close acquaintances had macro-PK ability.

INITIAL RESEARCH SESSIONS

Research sessions were held approximately 3 times each month, for 1.5 hours in the evening, in a video lab room of a social sciences building at the University of California, Irvine. The atmosphere of meetings was congenial yet purposeful. Participants sat around the room on chairs or couches, forming a rough circle about 15' in diameter. The first PK tasks always consisted of each person's attempting to bend rather lightweight cutlery via PK with touching. No effort was made to determine whether any PK was used to effect these bends, in a deliberately attempt to establish a psi conducive social context. After successful bends were being produced in this fashion by most participants, heavier weight cutlery and brass keys were introduced for similar attempts, as were more well-controlled PK tasks.

In about half the sessions a video recording system was used. Participants were taught how to operate the video camera, and were given opportunities to record live research footage and make suggestions to the others about how to improve the strength of evidence being gathered. By inviting participants to become co-experimenters as well as subjects, a greater cooperation, honesty, and interest seemed to be engendered.

Sometimes a participant could bend a spoon 45 seconds after having received it, and sometimes not for 45 minutes. Most initial bends seemed to occur after 5 or 10 minutes of attempts. Bending also seemed

to occur more quickly and easily in the 5 or 10 sessions that met on the night of the full moon; participants expected this, however, due to an initial voicing of the belief from early participants that the full moon enhanced PK attempts, and subsequent observations that successes seemed to increase on those nights, and this may have been the cause. Other researchers have noted a similar effect, nevertheless, observing correlations between telepathic success and the full moon (Krippner, 1975, 127-128), or, in rats, between precognitive hitting and the full moon (Treurniet & Craig, 1975).

Participants used various mental techniques for bending. One involved the imaging of the neck of the spoon softened, like putty, while trying to bend it, or visualizing the target bending or already bent. Another involved repeated mental issuance of a command to the spoon to bend. A third technique, and the most prevalent, was to conceive the idea or image of what was desired to take place - the paranormal softening and bending of the spoon - and then to begin rubbing and ignore the idea while engaging in conversation with others. This is similar to the 'letting go' method found to be used by British child PK metal benders (Cantor, 1979). Often conversations were interrupted as a speaker discovered a spoon ready to bend, at which point full attention was turned to the spoon. Bends also seemed to be performed semi-consciously, without a break in the flow of conversation during bending. Participants used different techniques at different times; however there was a tendency for the more experienced participants to use the 'letting go' method more often, suggesting that its use was predicated on having practiced imaging or concentration techniques earlier.

Very heavy gauge cutlery was bent and twisted by participants, some adolescents, in ways that the experimenter could not duplicate using the greatest of his physical strength. Brass keys were also bent, sometimes one-handedly by pushing the narrow section down with the thumb. The favorite target was silver-plated cutlery, primarily because silver was believed to be more susceptible to PK; silver-plated cutlery is physically more pliable than stainless steel, however.

The spoons, if touched in the area of bending immediately after having been bent, were usually only slightly warmer than room temperature, about as warm as if held in a closed fist for about 5 minutes. Generally the physically stronger participants produced the more difficulty and complex bends, with occasional exceptions,

suggesting that physical force was used by most participants for accomplishing most bends.

Perhaps the most interesting early result was the difference noted between what happened when the experimenter asked one of the physically stronger participants to bend cutlery using normal physical force, and then to attempt similar bends psychokinetically on identical cutlery. The exertion in physical tension in the muscles of hands, arms, and chest present in physical-only bends were absent in psychokinetic attempts, as were impressions left in the hand where the cutlery was held. Even more significant seemed to be what happened to the cutlery: in physically attempted bends either the item broke or microcracks or ruptures appeared on the surface in the area of the bend, quite visible to the naked eye. None of this was present on cutlery from psychokinetically-attempted bends, where the surface of the metal in the area of bends remained intact and shiny even on items bent to a degree considerably greater than the physically bent items. This same disparity showed between spoons bent physically by an outsider not claiming PK ability and cutlery bent in psychokinetic attempts by participants. It was this finding that prompted continuation of the research beyond the initial stage.

MORE CONTROLLED FINDINGS

Tasks were eventually introduced where either normal physical contact with the target or target area of a system was prohibited, such as by having participants attempt to bend a spoon without using the other hand or physical force to make it bend, or targets whose properties seemed to exclude the possibility of bending without fracture under ordinary, short term conditions were used, such as by using items of brittle metal as targets (as has been done in research by Hasted, 1976). Because of the rarity of success and frequent malfunctioning of the video tape system, only one of the results reported below, the causing of movement in a pendulum system from a distance, was recorded on videotape.

In one task participants were asked to hold a spoon in one hand and to concentrate upon but not touch the juncture between bowl and stem, to 'soften' it there via PK so that the bowl would droop over of its own weight. Two instances of 'droop-overs' were reported which were observed by other participants, but not by the experimenter. The bowls

were observed to droop over slowly toward the floor, over approximately a 30 second period, stopping after having bent downwards about 45 degrees.

A target system occasionally used involved a spoon clamped to a table by the stem, from the protuding bowl of which was hung a weight of approximately 5 lbs.. Participants were asked to concentrate on but not touch the spoon, to soften the juncture between bowl and stem through PK, so that once soft the weight would do the work of bending the free hand normally did. One instance of a bend with this target system was noted, while the experimenter was talking in an adjoining room with the participant who was focussing on the setup, which neither could see. The other participants were left unobserved in the room where the target was located, and none claimed to have touched it or approached it before, during, or after the bend took place, or claimed to see the spoon in the process of bending. The participants were engaged in attempts to bend their own targets and were socially conversing when the experimenter left and eventually re-entered the room. The participant with whom the experimenter was talking was observed to scrunch his face in mid-sentence, and to mention immediately thereafter that the experimenter should check the target. Upon doing so the spoon was discovered to be bent at a 90 degrees angle downwards, pointing straight to the floor. Such spoon-with-weight targets were never observed to bend spontaneously due to the 5 lbs. weight alone. Unfortunately, because the target was left uncontrolled with the other subjects the significance of this case is diminished. It should be noted that at no other time during the research was that participant's facial behavior observed to reoccur.

Ordinary hack saw blades were used as bending-with-touching targets because of their brittleness, which was believed to prevent their being permanently bendable without fracture to any significant degree (e.g. more than 10 degrees). The goal was to bend one in a U shape without fractures, hypothetically impossible using physical force under normal conditions and over a short period of time. One participant did this successfully on one occasion. In 3 - 4 discrete bends, which took place over a 3 minute interval and each of which was accomplished in about a second, he was able to bend a hack saw blade in a perfect U shape, no more than about 1" wide at any point. Many people, both participants and non-participants claiming no PK ability tried to accomplish a similar feat using physical strength and dexterity, and the blades were broken in every case without ever

having been permanently bent.

A target provided by an outside collaborator was constructed in the following manner. A small aluminum bar 3"x1/2"x1/8" was sprayed with a lacquer, and before this dried, small, extremely thin-shell, nickel-cadmium spheres about the size of salt graines were sprinkled over its surface, leaving only one small bare spot in the middle of one side. The bar was then sprayed again to adhere the spheres to it as well as possible. The spheres covered about 1/3 of the surface area of the bar, and were closely enough spaced that it was impossible to grip the bar without contacting a number of spheres. The spheres were brittle and fragile, so that exerting more than a very light force by touch, much less than that required to bend the bar, would crush them. If the bar was bent, examination of the spheres would give evidence of whether normal physical (contact) force was used. This target was placed on a table in front of a participant, who was asked to bend it without touching, and who then concentrated on it for about 20 minutes. At this point it spontaneously bent near the middle about 15 degrees, according to subsequent measurement. Another participant was observing at the time and did not notice any contact made with the target. The bend was reported to have taken place smoothly over about a 5 second interval. The experimenter and the guest experimenter were nearby and had periodically observed the participant concentrating and had also not observed any contact being made with the bar. Neither was observing when the bar reportedly bent. On examination a few spheres were found to have fallen off the bar, and this was thought to have been due to handling after the bend. A small fraction of the spheres were also found to have been broken, mostly on or near the ends, in such a way that this appeared to have occurred from handling and not an attempt to physically bend the bar. Though it was both the experimenter's and the guest experimenter's feeling that the bend was not fraudulently produced, because some spheres were broken the evidence of this case was considered to be inconclusive. This target design may well be useful for future macro-PK research, however, and is recommended as a fraud-sensitive test for macro-PK ability

NON-METAL TARGETS

During the research several targets and PK tasks were included that did not involve attempts to bend metal. These tasks were intermixed with the metal bending tasks that occurred in the latter part of each

session, and were included on an irregular basis. Several significant results were obtained with these non-metal targets.

A pendulum arrangement was constructed that contained two equal length 5" swinging pendulums in a closed plexiglas cube. One pendulum was selected as target and participants asked to make it swing in an indicated direction, without touching the cube or what it was positioned on, which was either a heavy table, a heavy trunk, a 3-legged plexiglas stand, or a heavy television stand with wheels. The second pendulum served as a control against using external physical force to make the target swing, so long as the target did not swing perpendicular to a line between the two pendulums which would imply application of a rotational force co-linear with the control pendulum.

On half a dozen occasions two participants were able to induce motion in the target pendulum while the control remained stationary, without having been observed to touch either the cube or its mooring. This was sometimes accomplished when both were attempting simultaneously to influence the system, and sometimes occurred when each was attempting alone. Motion was in the desired direction (not perpendicular to the line between the two pendulums) and resulted in a linear swing of up to 1/2", after a gradual acceleration from rest over about 5 or 10 minutes. The mooring and flooring underneath were solid and no motion could be induced by rocking the body back and forth in regular oscillations of various frequencies. Except for the hands and arms, neither of the two participants were observed to move their bodies during solo or tandem attempts. There were no earthquakes reported by the news during times of successful attempts. The occurrence was captured partially once on videotape, showing the acceleration from about 1/4" to 1/2" swing over a several minute period in the target pendulum while the control remained motionless. Usually about 5 minutes after the target had begun to swing the second pendulum would begin to swing slightly, probably because it was physically coupled to the target pendulum by being suspended from the same supporting member. PK successes were obtained with the pendulum system in about 6 of 30, 10-minute trials with bobs of plastic or glass faceted spheres about 1/2" in diameter. No successes were obtained after approximately 10, 10-minute trials each with bobs of wood spheres 1/2" in diameter, 3/8" metal bolts, or irregularly shaped 3/4" chunks of rose quarts stone. It was not clear why the latter pendulum bobs failed to be influenced.

Notable failures to obtain phenomena indicative of PK occurred when

participants were asked to bend-with-touching 1/8" and 1/4" thick plexiglas strips, to bend metal objects sealed in glass flasks, to concentrate at a distance on an ordinary liquid thermometer bulb to heat it, or to concentrate on a sheet of liquid cristal sensitive to temperature to heat it. A number of trials with these targets were attempted over several sessions and no evidence of influence observed, though with the latter targets temperature variations of under about 0.5 degrees would not have been observable.

OTHER RESULTS

Two other results from this research deserve mention.

First, when two or more participants attempted to influence the same target system at the same time, generally no enhancement of success seemed to occur. However for some, the setting of the research did seem to positively influence attempts; one of the successful pendulum movers never achieved a successful result when alone outside research sessions, while during sessions he achieved obvious individual success. Other participants noted no pronounced enhancement of their skills when with the group; indeed, some noted better success at home, in private, or when with other persons not involved with the research.

And second, there was one case of apparent steady development of ability of a participant, Dave. At first Dave was observed to be capable of only moderately difficult spoon bends, with what seemed a fair amount of physical exertion. He could not put any twists in the stem of a spoon, for example, a difficult PK bending feat. After several months he was observed to be able to put a twist, even two in the axis of a spoon in one or several separate PK efforts. Several months later he was observed to install three twists, as well as to be able to fold the bowl of a silverplated spoon in half back over on itself in a tight U. About a month later he finally was seen to fold the bowl of a stainless steel spoon over on itself for the first time. During later sessions of his 1.5 year involvement he apparently became able to exert influences without touching: this was one droop-over while holding a spoon stationary in the hand and the bend on the spoon-with-weight from the next room. Dave's apparent development continued until circumstances forced him to move from the area and discontinue participation, and it is possible his ability may have continued to improve had he remained. Only one other participant

showed an increase in ability, this involving becoming able to influence the pendulum system after about 8 months of participation in regular research sessions; this ability seemed to last for the remainder of the participant's involvement with the group.

DISCUSSION

The aim of the research was to determine whether or not participants were exerting macro-PK influences on target systems. Results, while not conclusive, indicate that on some occasions they may have been. The strongest evidence was judged to have come from: the effortlessness with which some physically impossible bends - for the experimenter - were observed to occur, some produced by young adolescents; the noted absence of fractures and ridges on items bent in psychokinetic attempts, as compared to items bent physically over approximately similar time intervals; the successful bending of a brittle hack saw blade in a U; the bending of a small aluminum bar without touching or producing effects designed to signal use of physical force; and the production of swinging action in one only of two adjacent pendulums sealed in a glass case. However, it is the feeling of the experimenter that when and if macro-PK was exhibited it was only infrequently. Similarly, a conservative estimate is that only 6 of the participants showed results suggestive that something paranormal might have taken place. Among these 6 were both adolescents and adults, males and females, the 6 were a representative sample of the entire group. Hoped-for videotape evidence of demonstrative events was never fully obtained; what the experimenter has is an edited videotape of several effortless appearing bending and twisting attempts and monitoring of some pendulum movements and accelerations.

Though certainly not conclusive, it should be remarked that under the moderately open conditions of the research, none of the participants was ever caught cheating or detected in a deliberate lie or act of fraud by the experimenter, other participants, or occasional guest observers.

A significant feature of this research is that it did seem to provide evidence pertaining to which mental techniques might be conducive for macro-PK. While visualization of a target softened, bending or already bent, and issuance of the mental command 'bend' were common mental strategies, the most extensively utilized was the

method of 'passive volition': not concentrating at the time of influence on the target or in any specific way, but just vaguely intending PK bending. That results often occurred with a lack of conscious technique and amidst casual conversation bears great similarity to the conditions under which the 'Phillips' group's PK effects reportedly occurred (Owen and Sparrow, 1976). Too, the idea that passive volition is a viable and important PK strategy has been discussed in Stanford's (1977) review of PK research, and has been mentioned in Braud's (1978) review of his own work. Based on the present study, this mental strategy would seem especially worthy of attention in future informal and formal experimental settings.

Future studies with persons claiming weaker forms of macro-PK ability may well benefit from group research sessions. It appeared that maintenance of the light-hearted though serious, casual (yet formal when necessary) atmosphere helped to elicit PK. These conditions are similar to those found to be used in some PK sitter groups, in which tables seem to be psychokinetically influenced (Batcheldor, 1966, 1979). Certainly the social interaction at meetings between persons holding similar beliefs both about the world and about themselves offered a reward that counteracted disappointment at the infrequency of occurrence of strongly evidential phenomena. The experimenter was relaxed, casual, friendly, and supportive during sessions as well as directive, in a distinct effort to break from the traditional aloofness associated with the role of experimental psychologist. The group developed a cohesiveness of purpose that allowed any individual's success on a task to be defined and felt as a success on the part of the group, and this too seemed to help foster a supportive atmosphere for PK.

The provision of a period in each session when bending attempts were not assessed for evidence of PK seemed to build morale and provide a momentum for later PK attempts. Participants could also get the feel of a few physical bends before trying psychokinetic ones if they wished. And once bends of any type began to occur they reported feeling like the ice had broken, that PK was happening in the room and that it therefore seemed more likely to happen for them. This supports the feeling of other experimenters that witnessing an event of apparent PK helps to stimulate later production of genuine PK (Brookes-Smith, 1973).

A future consideration for research of this type is the development of a macro-PK target that can be touched by subjects, to which some

ordinary physical force can be applied, and which will still yield evidence for paranormal influence if PK occurs. If such a target can be constructed that does not present a task so psychologically formidable that it is discouraging, then a fair assessment of the presence of weak macro-PK ability in the general population might be undertaken. More extensive tests might provide data on whether the ability can be taught or elicited in previously untalented persons; one participant in this research claimed he had 'taught' the PK-spoonbending skill to 23 of 25 persons he attempted to teach it to, though no attempt to assess the genuineness of this claim was made. If such a methodology can be found, it would allow the hypothesis to be tested that many persons possess some modicum of macro-PK ability, and that it is not just an anomalous skill possessed only by a few in the population

ABSTRACT

This paper presents observations from 2 years exploratory research on PK metal bending in a semi-formal group setting, with 13 persons from Southern California. Participants all claimed prior ability to bend cutlery or keys using either PK or a combination of PK and physical force. In research sessions they were gradually introduced to PK tasks that controlled for the use of physical force. From these tasks the strongest evidence for PK was judged to come from: the bending of a brittle hack saw blade in a tight U, without fracturing; the bending of a small aluminum bar without touching, or producing effects designed to signal use of physical force; and the production of swinging action in one only of two adjacent pendulums sealed in a glass case. Only 6 of the participants were felt to have produced strong evidence of PK. Passive volition appeared to be the predominant mental strategy. Research methods which emerged as psi-conducive in the group setting, and research targets potentially useful for future macro-PK work are discussed.

ACKNOWLEDGMENT

This paper was presented at the 23rd Annual Convention of the Parapsychological Association in Reykjavik, Iceland. My special thanks to Robert Morris, Stanley Krippner, and Ron Hawke for technical advice

and assistence.

REFERENCES

Batcheldor, K.J. 'Report on a case of table levitation and associated phenomena', J. S.P.R., 1966, 43, 339-356.

Batcheldor, K.J. 'PK in sitter groups', Psychoenergetic Systems, 1979, 3, 77-93.

Braud, W.G. 'Recent investigations of microdynamic psychokinesis, with special emphasis on the roles of feedback, effort, and awareness', E.J.P., 1978, 2, 137-162.

Brookes-Smith, C. 'Data-tape recorded experimental PK phenomena', J. S.P.R., 1973, 47, 69-89.

Cantor, R. 'The study: a summary', Human Dimensions, 1979, 6, 36-46.

Hasted, J.B. 'An experimental study of the validity of metal-bending phenomena', J. S.P.R., 1976, 48, 365-383.

Krippner, S. 'Song of the siren', New York, Harper & Row, 1975, 127-128.

Owen, I.M. & Sparrow, M.H. 'Conjuring up Phillip: An adventure in psychokinesis', New York, Harper & Row, 1976.

Stanford, R.G. 'Experimental psychokinesis', in Wolman, B. (ed) 'Handbook of parapsychology', New York, Van Nostrand Reinhold, 1977.

Treurniet, W.C. & Craig, J.G. 'Precognition as a function of environmental enrichment and time of the lunar month', in Morris, J.D. et.al. (eds) 'Research in parapsychology 1974', Metuchen, Scarecrow Press, 1975, 100-102.

Mark G. Shafer
McDonnell Laboratory for Psychical Research
Washington University
St. Louis, MO 63130
U.S.A.

THE VOICE ON TAPE PHENOMENA: LIMITATIONS AND POSSIBILITIES

Jurgen Keil
University of Tasmania

For a number of years popular books have been published in which the authors suggest and assert that tape recorders can be used to communicate with the dead (Jurgenson, 1967; Raudive, 1971; Schafer, 1978).

Even if one accepted some of the voices as paranormal the voice phenomena do not provide any support for the survival hypothesis beyond that which can be claimed on the basis of other parapsychological findings. Indeed compared with other material (see Gauld, 1977) the case for survival based on voice phenomena seems to be relatively weak. Furthermore in most if not all cases, it is extremely difficult to obtain any reasonable evidence for paranormality. The following discussion is therefore limited to items which have a bearing on this question.

A reasonable case has been made for psychic photography (Eisenbud, 1977). More recently investigations in the microstructure of objects subjected to PK (Hasted, 1978; Hasted & Robertson, 1980) provide a further suggestion that it is not unreasonable to expect that audio tapes may be changed in such a way that new and perhaps recognizable sounds appear on account of PK. Bender (1971) suggested that this was probably the case when a few of the voices associated with Jurgenson were investigated. Jurgenson has been most cooperative and as Bender's report indicates, a considerable amount of equipment was used in an

attempt to eliminate counter hypotheses. This resulted in a considerable reduction of the number of voices which were selected for further more detailed analyses. However, as Sotscheck (1971) pointed out speech diagrams although useful, do not provide conclusive evidence when the question of paranormality is considered (Sotscheck, 1979).

The strongest assertion were made by Raudive (1986; 1971; 1973; 1976) who spent years of listening to tapes and who registered and classified with impressive patience thousands of voices (7200 in 1971) until his death in 1974. Smith (1972) critically discussed these claims and came to the conclusion that they do not support the survival hypothesis and that little if any evidence for paranormality can be found. This author is in agreement with these conclusions but some of the reasons suggested by Smith (e.g. speaking in light trance) are perhaps somewhat improbable. This is understandable as Smith had no personal experience of Raudive.

Raudive was obsessed with his voices, willing to spend endless hours listening to them and discussing them and he had the obstinacy and a presumably superficial knowledge of science with which he managed to impress a number of people.

It is somewhat surprising that he had a good deal of support from many professional people who could be expected to appreciate the flaws in Raudives so called scientific tests. It is worthwhile to note that few if any parapsychologists (perhaps easily categorised as members of the PA) ever supported his claims.

The impressive care which Raudive lavished on his tapes and voices must not be mistaken for scientific inquiry. His listening test was a crude example of a clearly non-scientific test which nevertheless fooled many people. In this test Raudive announced the meaning of a particular sound pattern first, and then played this sound pattern repeatedly. If the listener agreed that he could hear the sound pattern as previously announced, Raudive listed this success of his suggestion as an independent verification of this particular voice. However, if a subject had doubts and expressed a low degree of confidence in Raudive's suggestion, he was classified as being unable to hear. Obviously this test has no scientific merit whatsoever. There is absolutely no doubt that sounds which have some similarity with speech and which are presented in short bursts only (say 2 to 10 words) can be interpreted in a variety of ways and most subjects

follow any suggestions quite readily.

Briefly, voices are said to have appeared under the following conditions:

A taperecorder with microphone attached was played, for instance, in order to record birth songs. Replaying the tape it was found that human voices appeared on the tape. It was assumed that the microphone could not have picked up the sound of normal human conversation.

This method is now used with the following main variations: The taperecorder with microphone runs (a) while nobody is present, and (b) to record a normal conversation with intentional pauses for a clearer manifestation of the paranormal voices.

It may be useful to designate the above variations as microphone conditions (1). Other methods of recording voices can be summarised as radio conditions (2). This includes the modulation or modification of an actual broadcast, i.e. the speaker seems to carry on but his words are changed in a paranormal way. This is also claimed for music. It is also assumed that on frequencies between stations, new words are formed apparently with the assistance of the energy available. In trying to find these paranormal radio voices normal receivers are used, usually in the medium frequency range. Some recordings have also been made with diodes which will pick up a wider frequency range. Small transmitters normally in use for test purposes have also been tried in association with a radio receiver. The transmitter is adjusted until it surpresses the normal radio broadcast on a particular frequency to which the radio is tuned.

It may finally be useful to include as an additional set of conditions, storage and replay conditions (3). Most of the investigators found additional voices when replaying a tape they had previously analysed. However, they assume that most, and perhaps even all, of the additional voices may have remained unnoticed during the earlier part of the analysis.

It is sometimes claimed that only through training over a long period will it be possible to hear and identify some of the weaker voices which have to compete with general noise on the tape, because they are only audible under extremely high amplification.

It is quite true that some improvement may be obtained through a

period of training, however an over-emphasis on this point has probably side-tracked the major problem whether evidence for the paranormality of the voices can be obtained. For this basic question only voices of considerable strength should be used which can be clearly heard by persons with normal hearing ability.

This author has clear evidence that Raudive who claimed to be a highly trained listener, made such substantial errors that it is highly likely at least in this case, that the repeated listening to faint sounds did not improve his ability to understand them but rather increased his tendency to interpret these sounds into familiar words, giving messages of a kind which he expected.

In 1970 Raudive was persuaded (with considerable difficulty) to make one of his original tapes available. For his 'tests' Raudive had copied only selected sections of this tape repeatedly and the original tape was obtained from which his test section had some of the clearest and loudest voices. In agreement with Raudive's wishes no copy of his original tape was made but during the period of a few days when it was available to this author (a native German speaker) it could be clearly established that the tape had a continuous German passage of about 37 words with some fading, making some of the words almost inaudible. A number of continuous words however, could be clearly identified and were recognized without exception (and of course independently and without prior prompting) by 8 native German speakers. Under similar conditions there was also unanimous agreement that the section was a continuous German passage. It was also established that the German words were spoken in such a way that the speaker was likely to have been a minister of religion or a person giving a religious talk. It was only after this conclusion had been reached that the date of the original tape recording was identified as Easter Sunday some years back. It was impossible to obtain detailed programmes about this day from German Radio stations but there seems little doubt that the recording (made under radio conditions) was part of a religious or similar broadcast by a German station. Even if one accepts some uncertainty as to the origin of the recording there is no doubt that Raudive who had a good knowledge of German but who is not a native German speaker, did not recognize the original passage as a continuous German one. Raudive used 5 languages to interpret the passage. He also catered for his apparent desire to be on good terms with famous authors of the past and cut up the original German word 'benoetigen' (clearly identified by all 8 German speakers) such that the 'noeti' part became Goethe. Raudive's messages in various languages had no

resemblance in content to the original German words which could be clearly identified.

It is perhaps not surprising that Raudive (not being a native German speaker and presumably with some hearing loss because of age) had more difficulties in recognizing the passage of words for what it was.

It is important to note though, that after Raudive had edited the original tape by presenting on a new tape only small cut out sections repeatedly, it was rather difficult to recognize the original words even if one knew them. Raudive himself saw the evidence for paranormality (and in his framework for survival) in the content of the messages which he recorded. As this example shows his interpretation had no resemblance with the original in a case which he himself had rated among the best and clearest.

This does not rule out the possibility that Raudive had recorded some paranormal voices but on the basis of the above example it seems fair to conclude that there is no hope of extracting them from among his enormous amount of tape material with any degree of assurance. In an independent investigation Ellis (1978) came to a similar conclusion although he had perhaps started off with greater hopes of finding support for Raudive's claims.

Even if the most publicized case for voice phenomena is extremely weak it does not necessarily follow that others may not succeed where Raudive has failed.

It is necessary though to emphasize that voice recordings include a vast variety of possible errors, some of them decidedly self deceptive.

Special difficulties which may lead to errors are the following:

(1) The voices speak in languages not known or not well known to the listener. On the basis of the experience gained so far, it seems quite pointless to deal with these examples. Of course they may be considered by others who are native speakers of the language concerned.

(2) The voices speak with a peculiar rhythm which is supposed to identify them as paranormal. There are numerous ways through which normal radio voices may sound like spoken with an unusual rhythm,

simply through some oscillation of the radio wave which is received. There is further evidence that radio voices which were identified as normal radio voices with a high degree of confidence and which did not show any of the rhythmic pattern, were not recognized as normal. By unusual editing (e.g. cutting out sections from a single word, see the Goethe example above) the normal speech rhythm may be disturbed and the impression of a peculiar rhythm may be formed.

(3) Raudive claims that in order to maintain this rhythm the normal words are shortened or lengthened by changing the number of syllables in a word. Obviously this allows for an increase in faulty interpretation, i.e. the probability is much higher that the so-called changed word is really a normal word in a different language or perhaps two words.

(4) During normal fading of radio programmes this process is a gradual one. In some recordings relatively loud sections appear suddenly and disappear suddenly. However it was demonstrated in the Institut at Freiburg by Mr. Lemke (Munich) that circumstances may occur where through the introduction (switching on) of an additional station, the wavelength of another station may alter instantaneously. Sudden appearance or disappearance can therefore not be taken as evidence for paranormality.

(5) It is claimed that the voices speak during one sentence with different languages. In the extreme case five words may be spoken in five different languages. This claim is very difficult to verify as should be clear from problem no. 1. Most listeners cannot claim a really good knowledge of five languages and probably native speaking is highly desirable. The assumption that many languages are used, obviously makes self deception much easier. With some knowledge of some languages it is always possible to find a suitable meaning in some sequence of words which are only half audible and which may well all be spoken in a different language from those assumed to be used. Again the investigators of voices do not use this criterion consistently, i.e. a sentence may appear in one language (if the contents seem to be right) or in two or in more as the case demands. If it can be claimed that a few examples exist where in a sequence of words a clear change from one to another language occurs, it is still possible that this might have happened in a modern radio play or interview, but also through shortwave telephone transmission which under some circumstances may intermittently and in a mixed up way intrude into radio programmes (Lemke, 1970).

In summary, it may be said that as far as the radio conditions are concerned it is extremely difficult to exclude the possibility that a normal radio or similar transmission is mis-interpreted as paranormal.

It seems hardly worthwhile for a parapsychologist to start an investigation unless extremely clear and really extraordinary passages are presented. There will be some subjective differences as to what constitutes an extraordinary passage; it is the opinion of this author that there are few, if any, recordings that were done under radio conditions which would qualify.

Under radio conditions one research possibility exists to some degree, that is, if modulations of normal radio programmes are claimed to occur through paranormal processes. If it was possible with the assistance of a station, to obtain a copy of pre-recorded broadcast and, if this broadcast was monitored by a number of people located at various distances around the voice investigator as well as by the voice investigator himself, then a clear change of a broadcast would amount to an interesting phenomenon which is perhaps not yet convincing but which has possibilities to be controlled further. Longer recordings of the broadcast allowing for comparison between different receivers for periods before and after the voice phenomenon is claimed to take place, seem desirable.

(6) A further error may be introduced if in a microphone recording during which some conversation occurs, unconscious whispering or speaking takes place. There is little doubt that through the introduction of partly directional microphones some cases of unconscious speaking could be detected with virtual certainty which under different recording conditions most likely would have been assumed to be of paranormal origin.

The technique of directional microphones is relatively simple and only requires one independent track and one independent microphone per participant.

In addition there should be at least one further centrally placed microphone connected to one further independent tape track. A normal microphone can give sufficient level of directionality by placing a narrow conical cardboard tube around it into which the participant speaks.

If anything unusual appears on tape later it can be clearly

established who spoke during the period (as well as before and after) and whether the ostensible paranormal sound originated from one particular microphone or not. If this is the case the probability is high that the unconscious speaker is found. It should perhaps be emphasized that this unconscious talk goes on usually more frequently than is generally realized and even serious investigators who are determined not to make any sound are not free of it. It seems to be particularly prevalent if one participant wishes to say something but is prevented from doing so by the continuation of speech by someone else.

Spontaneous agreement with a remark of a speaker may also result in audible exclamations. It is important to note that serious parapsychologists were largely unaware of this. One of the most experienced investigators of the Freiburg team said spontaneously with the fullest conviction that particular sound phenomena which were suspected as originated from a particular speaker could not possibly be the voice of that person. Yet it was subsequently established that it was that person who had spoken without awareness. This example should clearly show that it is necessary to be extremely careful in order to avoid this kind of error.

A simple test under microphone conditions which can be clearly evaluated may be to divide 20 minute recording sessions into 10 two minute trial periods during which according to a random number table approximately half are supposed to remain silent while the other half is supposed to receive as many voices as possible. Independent judges could later estimate which trial periods according to some specified criterion are more or less silent. From a psychological point of view, this framework may not be suitable for some people but with due regard to the particular claims and idiosyncracies of particular people it will be necessary eventually to test the claims in such a way that both support or lack of it for the psi hypothesis can be stated in probability terms.

Altogether this seems a fairly negative account. It is written with the hope that it may help parapsychologists who are confronted with these claims for the first time. It may also help to start research with a better chance of tackling the problem involved.

There is also a very positive aspect in the numerous possibilities of self-deception to which the voice investigator is subjected. A medium trying to move an object through PK and who is actually able to

do it once in 500 trials can see in 499 trials that nothing is happening and he is likely to get discouraged. The voice investigator believes he is getting a rich harvest of voices practically all the time and this may actually ensure that a few genuine ones appear among many errors.

If these genuine phenomena are there, this situation which is rewarding for the medium, may actually persist long enough to allow parapsychologists to extract the genuine cases with sufficient evidence.

REFERENCES

Bender, H. 'Zur Analyse aussergewohnlicher Stimmphenomene auf Tonband', Zeitschrift fur Parapsychologie, 1970, 12, 226-238.

Eisenbud, J. 'Paranormal photography', in Wolman, B. (ed) 'Handbook of parapsychology', New York, Van Nostrand Reinhold, 1977.

Ellis, D.J. 'The mediumship of the tape recorder', Ferwood Nightingale's West Chillington Pulborough West Sussex, D.J. Ellis, 1978.

Gauld, A. 'Discarnate survival', in Wolman, B. (Ed) 'Handbook of parapsychology', New York, Van Nostrand Reinhold, 1977.

Hasted, J.B. 'Merkmale paranormaler Metallbiege-Phanomene', Zeitschrift fur Parapsychologie, 1978, 20, 173-184.

Hasted, J.B. & Robertson, D. 'Paranormal action on metal and its surroundings', J. S.P.R., 1980, 50, 378-398.

Jurgenson, F. 'Sprechfunk mit Verstorbenen', Freiburg, Hermann Bauer, 1967.

Lemke, K. Private communication with Keil, 1970.

Raudive, K. 'Unhorbares wird horbar', Remagen, Reichl, 1968.

Raudive, K. 'Breakthrough', New York, Taplinger, 1971.

Raudive, K. 'Uberleben wir den Tod?', Remagen, Reichl, 1973.

Raudive, K. 'Der Fall Wellensittich. Untersuchungsbericht zur Frage der Medialitat bei sprechenden Vogeln', Remagen, Reichl, 1976.

Schafer, H. 'Stimmen aus einer anderen Welt', Freiburg, Hermann Bauer, 1978.

Smith, E.L. 'The Raudive voices - objective or subjective? A discussion', J. S.P.R., 1972, 46, 192-200.

Sotscheck, J. 'Uber Moglichkeiten der Erkennung von Sprachlauten', Zeitschrift fur Parapsychologie, 1970, 12, 239-254.

Sotscheck, J. 'Uber tauschungsmoglichkeiten bei der Beurteilung von Ergebnissen aus Abhorversuchen von Tonband-Sprachaufnahmen', Zeitschrift fur Parapsychologie, 1979, 21, 41-51.

DEVELOPMENT OF THE SYSTEM THEORETIC APPROACH TO PSYCHOKINESIS

Walter v. Lucadou and Klaus Kornwachs
Institut fur Grenzgebiete der Psychologie
und Psychohygiene
Freiburg
West Germany

1 - Why do we need a system theoretical approach to parapsychology

Cybernetic or general system theory has become a very successful instrument in almost every field of science during the last decades. This has become obvious especially in the field of biology and in the social sciences. Its application in technology, for instance in computer science and economic sciences, has lead to great success which seems to be comparable to the industrial revolution of the nineteenth century.

The question of why it has been so successful will lead us directly to the answer to our question of why it seems useful to apply system theory to parapsychology. It is often maintained that the system theoretical approach to parapsychology is a physicalistic one, but that is not quite right; it is neither a neurophysiological nor a psychological approach. System theory does not deal with the substratum such as physics and physiology do, but with special structures in the field of concern. Therefore it is often the case that a physical system of classical mechanics, an electromagnetical system, an economic system or a psychological system are described by exactly the same differential equation and therefore are from the

point of view of system theory just the same. This is a very important feature which we should keep in mind in further considerations. So we can say that system theory has a kind of metatheoretical character with regard to the normal field of sciences such as physics, chemistry, and so on. This means however that the terms used in system theory or cybernetics (for example the notion of information) cannot be reduced to physics as W. Buchel (1975) has shown but on the other side it is not necessary to introduce a dualistic world view as it is often done to solve the mind-body problem (see for instance J. Eccles (1977) model of a dualistic world).

Here I see the main advantage of applying system theory to parapsychology for it becomes possible to describe physical and psychological features in a joint theory without using incommensurable terms of dualistic worlds which may lead to great problems if one tries to falsify such dualistic theories in the sense Popper demands that it be done.

2 - Problems linked with the socalled Observational Theories

What is mentioned in the section above shows us the main problem of the observational theories (for instance Walker's theory (Millar, 1978)). In science most observational theories are purely physical theories, they do not explain how psychological variables can be fitted into the framework of the theory. It has not yet been shown that psychology is reducable to physics for the reasons mentioned by W. Buchel (1975). I do not believe that it is at all possible. From our empirical research in parapsychology we know how important psychological boundary condition are for any paranormal effect. It is quite reasonable to suggest that on the other hand psychological conditions are at least correlated with brain-states, therefore, one has to put the question: what is the interface between the brain and the physical quantum process within a PK effect for instance. I think it is not sufficient when Walker states that in the synaptic cleft of a neuron there can occur quantum processes which are linked with an external non-local wave function of the PK process. Since he puts hidden variables of the wave function on a level with consciousness his view is strictly speaking a dualistic one. If one wants to avoid this dualistic view one has to ask what is the mechanism controlling the quantum processes in the synaptic cleft that is dependent on psychological conditions.

Another problem is that most observational theories (see Walker, 1975; Mattuck, 1978; Sarfatti, 1974) are related to a special interpretation of quantum theory. This, I believe, is one of the basic reasons why J.A. Wheeler (1979) has aimed his criticism at parapsychology rejecting a special interpretation of the measurement problem used by Sarfatti and others. Especially these subjectivistic interpretations lead to the curious question of whether one has to attach consciousness to a quantum mechanical measurement apparatus. I would not say that the interpretation of quantum theory is a useless question, but since the output of a quantum mechanical experiment does not depend on the interpretation of the formalism but on the formalism itself it would be better to have a theory which is invariant in regard to the interpretation of quantum theory. As we can see later within our system theoretical and quantum theoretical approach the question of interpreting quantum theory is irrelevant insofar as every interpretation that is discussed is applicable.

One certainly could put forward the objection that such an approach is too restrictive. Because while dealing with psi effects which are independent of time and space one has also to deal with the basic principles of our understanding of space and time which is certainly linked with the problem of interpreting quantum theory. Therefore it is often speculated that we can solve our problems by applying elementary particle physics or cosmological considerations to parapsychology. The only thing I can say to these attempts is that to my mind it is not only a question of the scientific economy principle but a question of the amount of energy which is involved in human metabolism. This idea was discussed by J. Taylor (1978) before he withdrew from parapsychology. But it does not imply, as Taylor (1978) thought, a restriction of the involved interaction to electromagnetism.

On the other hand I do not really see how a possible electromagnetic influence on the brain's psi ability can be incorporated into the observational theories especially in Hasted's theory of a multiple universe, even if this electromagnetic influence should be established as P. Maddock (1978) shows in his work. Another problem is how to understand the rather specific analogy between alpha-rhythm biofeedback and PK effort as J. Beloff (1979) has pointed out.

Synchronistic events present more of a problem. Many parapsychologists, such as J. Eisenbud (1979), believe that synchronistic events can be reduced to ESP or PK. But H. Bender (1977)

has put forward many examples which cannot be 'explained' in this way. He calls them 'meaningful coincidences', and we will see that the word 'meaning' is a key-word used in our theory.

Finally let us discuss the very important divergence problem. Since in the observational theory of Walker and Schmidt (1977) the outcome of a PK experiment is dependent on the psi abilities of the observer (this might be the subject, the experimenter, or even another person) the experiment can never come to an end, for Schmidt has shown experimentally that PK is time-independent and a PK influence on already published experiments seems to be possible in the context of the theory. Seeing this obviously unreasonable feature of the theory both Schmidt (1978) and Walker tried to show that the effect will be cancelled out by the positive and negative psi sources of later observers. But since we do not know anything about the psi-abilities of our descendants I am rather unsatisfied with this solution. On the other hand within a system theoretical approach this problem could be solved.

3 - Properties of complex systems

Most parapsychologists would agree that psi abilities are coupled with brain or with neurophysiological activity. Now, what's the difference between a brain and other matter, for instance a piece of wood? It cannot be the amount of molecules, it cannot be the chemical composition, it cannot be the very complicated structure, for a piece of wood is complicated too. It cannot even be the fact that it is a living system, for neurons are the only cells within the body which cannot be regenerated. From a point of view of general system theory the important difference between a brain and other matter is the brain's enormous degree of organization or, in other words, it's enormous complexity.

Thus we must primarily explain what the notion of complexity means. But before doing this, a few remarks on the term system. A system can be defined (Klir, 1969) by the set of the elements and the set of relations between these elements. The totality of these sets is called structure. The elements have properties which can vary in time. The set of the properties during a time 'dt' (conveniently short) is called a state.

Time dependent systems can be defined by a state-to-state transition table which is equivalent to a Turing machine. A system of such kind we call a system of the first type. From a structural point of view the existence of relations between elements can be classified by writing down a density of structure which contains the actual connections, divided by the number of possible connections, depending on the number of elements. It can be imagined that there are some areas in the system which have a denser structure than other ones. In this way it is possible to define a hierarchical substructure which leads to different hierarchical levels. Starting from this point one can derive a measure of structural complexity. As it is shown elsewhere (Kornwachs & Lucadou, 1975, 1976, 1978), from a functional point of view one can construct in a similar way a measure of functional complexity. The important feature of this measurement of complexity is, that it rather more depends on the number of hierarchical levels which the system has than on its number of elements. This is an important feature of what we call organization. Now it may be possible that there are systems which are able to alter their own structure in time, we call them systems of the second type.

If we now regard these very general systems where nothing is said about the physical substratum or the special kind of interaction between the elements we can even find a lot of interesting questions in respect to the behaviour of such complex systems. The first question is whether we can observe the system. This means that it must be possible to interact with it and, as we shall see, this question is not a trivial one. The second question is whether the system is stable at least for some time. The third question is whether we can influence the system.

Even if we cannot follow these very interesting questions too far it should be mentioned that even systems of a rather low degree of complexity can show perplexing qualities. So S. Kaufmann (1969), M.R. Gardner and W.R. Ashby (1970) and others have performed investigations on systems with elements which are connected at random and which show interesting behaviour patterns in the realm of stability but on the other hand show a rather abrupt breakdown of stability dependent on connectance. This behaviour reminds one of properties of physical systems far from the thermodynamic equilibrium, which are theoretically described by the theories of I. Prigogine (1978). He has pointed out that such systems can produce a high degree of regularity over a large distance of time and space. The description of these coherent space-time structures which emerge from 'giant fluctuations'

shows clearly that the conception of direct microscopic interaction has to be replaced by the concept of correlation. In this context I want also to refer to the most exciting research done by F.A. Popp (1979) and H. Haken (1977) who investigated long range coherent radiation within living systems which exhibit 'collective phenomena' which resemble a laser mechanism. Nevertheless these systems are not very complex in comparison to the brain.

If we now look at complex systems of the second type, which can alter their structure, we see that their degree of complexity may not be constant in time. Then it can be shown that, from a logical point of view, the terms of description contain predicates of predicates expressing properties of properties. In other words, one must use a metalingual level in order to describe the behaviours of the system, and this means the use of predicate calculus of the second step. Using the famous theorem of Godel (1931) which proposes that such a description is not complete, one can recognize that complex physical systems of the second type can always produce behaviour patterns which cannot be described within a given theoretical system.

This leads to a principal lack of information about the physical system so that from a certain degree of complexity it can only be described as a stochastical system. Or, in other words, it is no more a deterministic one. Since this lack of information is of principal nature and cannot be removed as a consequence of Godel's theorem, the stochastical behaviour of an element of the physical system cannot be distinguished from quantum mechanical behaviour. This does not mean that we try to derive quantum theory from Godel's theorem. It has already become clear that this is an unsuccessful approach.

In our contribution to the Parascience Conference in London in 1977 (Lucadou & Kornwachs, 1977) we stressed the consequence of this assumption for parapsychology. It leads to the statement that in complex physical systems which can alter their degree of complexity the Schrodinger equation will become non-linear and time dependent with respect to the macroscopic system theoretical state of the system. But this is exactly the unknown interface between the cybernetic description of the brain and the mechanism of the collapse of the wave function in observational theories (see Mattuck, 1978). The advantage, however, is that we need not make any additional assumption on the collapse of the wave function or on the interpretation of quantum theory since the influence of the system can in principal be inserted explicitly in the Schrodinger equation.

Needless to say, in practice we are far from this point.

Finally there remains the question whether one can construct or simulate a system of such properties since it would not be very good from a point of view of philosophy of science if the brain was the only example because this approach has been taken in order to obtain a model of some of the brains special features.

I believe that we will soon find such models if the density of integration in microprocessor-chips is increased. Surely microprocessors are constructed under the option of discrete Boolian logic. But if the density of integration becomes higher and higher one will reach the point where the function of a single transistor element will depend on the function of its surroundings which normally should not influence it. The number of internal states far exceeds the possible output states of, for instance, fourty output pins, and by combinatorial means it can be proven that for the realization of a distinct output function there can be several internal states. This is the reason why greater expenditure is needed to test a microprocessor for reliability (Kewney, 1978). The result is that such microprocessors may to a certain extent show indeterministic behaviour. But as it was discussed above an indeterministic quality may not necessarily lead to chaos; what we would rather expect are coherent space-time structures like I. Prigogine found. In the case of complex systems we should call them 'collective' phenomena.

4 - The human brain under the aspect of system theory

It is clear that we cannot as yet give a sufficient system theoretical description of the human brain, for our knowledge of neurophysiological details is far to low. But since we know special features of important elements of the brain namely the properties of neurons we can try to develop a simple model of the brain and compare it with the empirical facts. One prominent quality of neurons is that they behave like a threshold-switch. This model of a neuron is called the McCulloch cell (Arbib, 1964). We will now take properties of networks built up by such McCulloch-cells into consideration. In the last years this field called threshold-logic has become a special subject of cybernetics with wide-spread literature (Hurst, 1971) and practical applications. Here we can discuss only some elementary features.

A McCulloch cell or threshold module has multiple inputs and one output. The behaviour of the threshold module is characterized by the number of its inputs and the threshold of the module. If the amount of simultaneous input pulses overcomes the value of the threshold then and only then the output will be fired. If one now connects the output of such a module to an input of some other modules, and so on, one gets a modular network.

With the tool of threshold logic it can be proven, that it is possible to construct even very complicated Boolian structures as they are realized for instance in normal computers with less elements than in normal switching systems. Furthermore it is possible to construct those modular nets which have the important property of stabilizing a special logical function so that it becomes independent from the value of the threshold over a wide range.

On the other hand it can be shown that the logical function of such a network can change very drastically only by lowering or raising the threshold. Thus, if one allows that the threshold level of the modules can be altered by the system itself we will get just a system of the second type mentioned above. Then the functional structure of the system can be changed enormously without any change on the hardware side. This means that the functional degree of complexity becomes a dynamic variable of the system while the structural degree of complexity remains unchanged. But this is analogous to what we observe in neurology: the hardware structure of the human brain is developed in the early days of childhood and remains nearly unchanged during the whole of life, whereas the intellectual evolution can stay in progress up to old age.

Now it becomes clear why the human mind on the one hand can exert creative processes and free will as we shall see below but on the other side may not descend precipitously into the chaos of instability as one would expect from the enormous connectance of the brain. The human mind is able to stabilize itself and to perform normal tasks as are necessary for every day life. This seems to me as wonderful as a PK effect.

Until now it is not clear enough why we maintain that determinism breaks down to some extent in the realm of the mind-brain system. Therefore we must elucidate the relationship of language to consciousness. As B. Honegger (1979) has shown in her paper presented at the 1979 PA Convention one is forced from a neurological point of

view to regard consciousness as being closely connected to the neurological substrate for language which is normally lateralized in the left hemisphere of the human brain. However, one should not regard language and consciousness as being equivalent, but under certain circumstances language can be a hint at consciousness. Certainly a computer language is not rich enough since it is performed on the basis of Boolian algabra, but the fact that natural languages cannot be completely formalized may be such a hint. A computer can never use or utter ambiguous statements, but the human biocomputer, as J. Lilly (1967) calls the brain, can do so. This fact shows that the system cannot be a completely deterministic one. Many efforts have been made to solve the problem of free will, and many philosophers have maintained that it does not really exist. The question is of great importance in parapsychology in regard to precognition and backward causation, but in dealing with only one aspect of this interesting question I would like to refer to the very lucid papers of J. Beloff given at the PF-Conferences in 1976 and 1978 about 'Backward Causation' and 'Voluntary Movement, Biofeedback Control and PK'. It is often maintained that free will emerges from a kind of random generator inside the brain. The eminent German physicist P. Jordan (1947) imagines that quantum processes in the brain could be amplified and lead to actions or thoughts which are interpreted as free will. J. Eccles (Popper & Eccles, 1977) follows the same line in interpreting the experiments of Kornhuber, though his view is a dualistic one, taking into account that "hundreds of thousands of neurons would be modified as the result of an 'influence' that initially caused the discharge of merely one neuron". However I believe that the subjective feeling of most people in deciding voluntarily is not that the decision has come about by chance, but by the whole background of their thoughts and feelings and so on. Nevertheless the decision needs not be a purely deterministic one, but it is only comprehensible in the context of the system as a whole. This is what we would expect the nondeterministic features of a complex system to be.

Finally we should resume the problem of what we have called 'collective phenomena'. It is a well known fact, that conscious notions and feelings cannot be localized at just one point at the cerebral cortex but are spread out over a certain area. For our consideration it is not necessary to decide whether we like to agree with K.H. Pribram's (1979) idea of holographic storage or with J. Eccles conception of modules. The important statement seems to be that it is obviously a property of the brain that the state of a single isolated element of the cortex is no more responsible for the

collective states embedded in the system in such a strict deterministic matter as it is in our modern computers. This shows from another point of view that with a complex system it is no more possible to describe the system by the properties of the elements alone, without using metalingual terms or a stochastic description.

5 - A modified conception of information and its importance for parapsychology

The concept of information is often used in parapsychology in the sense that a person receives information by telepathy or clairvoyance or even precognition. J.B. Rhine's classical performance of ESP experiments is constructed upon the tacit model of a subject as a transmitter or a receiver. Moreover in Walker's theory of PK a certain amount of information must be processed by consciousness to collapse the wave function.

But from the point of view of system theory this concept of information requires some attentive consideration. The usual concept of information in the sense of Shannon requires a transmitter, a channel, and a receiver. The measure of information, called transinformation, which is processed in the channel is a function of the probability of the symbols of the transmitter and the probability of the symbols of the receiver and finally of the probability of the pairs of symbols that occur.

Information, however, which has been 'understood', must induce a certain change in the receiving system - otherwise this information is not observable for the system. The effect the information has however cannot be described by the information concept used in the sense above, since the received information alters the conditions of the receiving system. In other words: The probabilities of expectation are no more constant. That means that the reception of information can be formally called a process of learning. The system reacts, its behaviour is altered. Therefore the information transfer can no longer be defined as a property of the channel, it must be interpreted in terms of the change of the structure of the system.

Such a concept of information which includes its effect on the system or in other words its 'action' has been put forward by von Weizsacker (1974). He shows that such a conception of information

which he calls 'pragmatic information' must be considered as a kind of product of the concept 'novelty' and the concept 'confirmation'. These two concepts seem to show complementary structure. If you read a Chinese newspaper and you cannot understand the chinese language then you will not get any 'pragmatic information'. On the other hand if you read yesterdays newspaper which you already know then you will not get any 'pragmatic information' either.

In regard to a system with consciousness one can identify the concept of 'pragmatic information' with the notion of 'meaning'. This however plays an important role in parapsychology. Especially synchronistic events are regarded as meaningful. Spontaneous cases of ESP and PK often exhibit a special 'meaning' in particular situations. H. Bender calls this the 'emotional field'. In respect to this we will now concentrate on one of the most important features of paranormal events.

They exhibit a similar complementary structure too as we have found with the concept of pragmatic information i.e. novelty and confirmation. It is often said the paranormal events have an elusive character. Especially in RSPK cases the phenomena seem to be 'observer-shy' (Bender, 1974). W. Gauger (1977) calls this the 'gap-phenomenon'. Even in the experimental situation in the laboratory we observe this feature. In his paper mentioned above J. Beloff (1977) shows that there are a lot of similarities between PK-effort and alpha-rhythm biofeedback control. The paradoxical situation that any conscious effort to produce an effect will make it vanish, reveals the asserted complementary structure between accurate control and the potentiality of an effect.

To elucidate the circumstances let us now regard a kind of 'Gedankenexperiment' which we will actually perform at Freiburg as soon as possible. Let us begin with a PK-experiment with pre-recorded targets (Schmidt, 1976). As it is known with this experiment the random sequence is not presented directly to the subject like in normal PK experiments but is stored in a memory before the very beginning of the experiment. By multiple storage it is asserted that the contents of the memory cannot be affected directly by PK. The important point is that no one, neither the subject nor the experimenter can know whether a pre-recorded target is presented or not.

There now arises the interesting question of whether an inspection

of the pre-recorded data before the test session will alter the outcome.

First let us suppose that it will not do so. This would mean however that we would find a deviation from chance of the frequency of the 'hits'. We could now decide to regard the inspected runs as control runs, which will not be presented to the subject. This however would lead to the paradoxical situation that some control runs show that the random generator is working correctly, whilst others show that it is malfunctioning. From this wellknown 'intervention paradox' we see logically that our supposition must be wrong. Therefore we would expect that such prerecorded runs which will be inspected before show no significant deviation.

But now let us suppose that the person who inspects the prerecorded runs neither knows the expectation of pure chance, nor does he even know what he is really doing. In short he does not know the 'meaning' of the situation. In this case the outcome of the experiment would still be left open. In this case the person could not make his decision upon the knowledge of the system but only by pure chance. Under such conditions there arises a lot of interesting questions.

What could be expected if the person does not remember the result of his efforts. What could be expected if he remembers the data but does not know its meaning and tells it to a person who is well informed of the meaning. These questions are still open, but we see a large field of new experimental possibilities.

An important advantage of this experimental approach is that it has become possible to connect a psychological question directly to the experimental setup as a kind of on-line variable. At our institute at Freiburg we are just beginning with the preparations of such experiments. What we would expect, is that the more the person who inspects the prerecorded runs knows about the whole situation the more the PK effect must decline. Knowledge of the system is just the same as the meaning of the situation or what we have called 'pragmatic information'. So we can say that the more we know about the system the less it can percipitate a paranormal effect, or in other words the more we achieve a deterministic description of the system by information of the system the less it can show an indeterministic behaviour. More generally speaking this reveals the complementary structures of the notions of behaviour and determinism, which in the realm of classical ontology does not have this quality.

We now formulate the following, rather speculative, uncertainty principle for complex systems:

> The behaviour of a system versus its deterministic description is limited by the least action of pragmatic information

This supposed least action of pragmatic information we want to determine in our experiments. This means a quantitative measurement. Starting from such an uncertainty principle we get a new understanding of paranormal phenomena.

The main difficulty of explaining paranormal phenomena has been their alleged quality of being unlimited in time and space. This seemed to contrast with their rare presence and their elusive character. But from our point of view paranormal phenomena are no more unlimited but can only be spread out to a certain extent and are restricted by the uncertainty principle. At this point it becomes obvious that there can no more be a 'divergence problem' like in usual observational theories. From this we would expect that large scale psi phenomena can only occur in situations were something remains unclear, where one cannot be sure about all circumstances, and this is exactly what we observe and what was called the gap phenomenon. This seems the basic reason why poltergeist phenomena are so elusive and the controversy about them will never come to an end.

Let me hasten to point out that this does not mean that the deficiencies of the empirical evidence for psi phenomena should be justified from a higher metaperspective with a kind of immunizing strategy which can never be refuted by empirical data. On the contrary, there exists an analogy to quantum theory where the uncertainty principle does not imply that precise quantitative measurements and predictions cannot be made.

Now in concluding our considerations we want to discuss the question of applicability to parapsychology. Here I only want to refer to direct application of psi effects as information transfer by telepathy or predictions by precognition. In this field we have seen that we must at least be careful with the application of the usual term of information. I will demonstrate this with the problem of precognition. The only thing we can find in an experimental set up is a correlation between the calls and the targets. Let us suppose that this correlation is really obvious. In this context there is often mentioned the famous Einstein-Podolsky-Rosen-Effect (Costa de

Beauregard, 1976). In an analogous way it was often speculated whether one can transmit information faster than light by this effect. Here one has a correlation between two outputs too. In terms of usual quantum theory the correlation occurs instantaneously. Now the question arises: can we transmit information with such devices.

In both cases there is a positive correlation but the discovery of this correlation is only possible post hoc and by means of normal physical information processing. And this is what we observe in parapsychology, a paranormal effect can be discovered only post hoc and until now there is no possibility of prediction or guaranty of success. Most parapsychologists are of the opinion that with further investigations it will become possible to make such predictions but this is a belief which is not supported by empirical data. From our point of view, which I tried to explain in this paper, we can only establish a correlation 'post hoc' for if we could do this beforehand we would fix the system to a deterministic one so that it can no more elicit any psi effect. On the other hand this does not mean that we cannot use this correlation to some extent in a practical way, but I believe that nowadays we are still far from this starting point.

ACKNOWLEDGMENT

This paper was presented at the Parascience Conference, London, 1979.

REFERENCES

Arbib, M.A. 'Brains, machines and mathematics', New York, McGraw-Hill, 1964.

Balanowski, E. & Taylor, J.G. 'Can electromagnetism account for extrasensory perception?', Nature, 276, 1978, 64-67.

Beloff, J. 'Backward causation', in Shapin, B. and Colin, L. (eds) 'The philosophy of parapsychology', New York, Parapsychology Foundation, 1977.

Beloff, J. 'Voluntary movement biofeedback control and PK', in Shapin, B. and Colin, L. (eds) 'Brain, mind and parapsychology', New York, Parapsychology Foundation, 1979.

Bender, H. 'Modern poltergeist research - A plea for an unprejudiced approach', in Beloff, J. (ed) 'New directions in parapsychology', London, Elek Science, 1974.

Bender, H. 'Meaningful coincidences in the light of Jung-Pauli's theory of synchronicity', in Shapin, B. and Colin, L. (eds) 'The philosophy of parapsychology', New York, Parapsychology Foundation, 1977.

Buchel, W. 'Naturphilosophische Uberlegungen zum Leib-Seele-Problem', Philosophia Naturalis, 14, 1975.

Costa de Beauregard, O. 'Time symmetry and interpretation of quantum mechanics', Foundations of Physics, 6, 1976.

Eisenbud, J. 'Synchronicity, psychodynamics and psi', PA Convention, 1979.

Gardner, M.R. & Ashby, W.R. 'Connectance of large dynamic (cybernetic) systems: Critical values for stability', Nature, 228, 1970.

Gardner, M. & Wheeler, J.A. 'Psi Quacks', The New York Review of Books, XXVI, 8, 1979.

Gauger, W. 'Das Lucken-Phanomen', Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie, 19, 1977.

Godel, W. 'Formal unentscheidbare Satze in der principia mathematica', Monatshefte fur Mathematik und Physik, 38, 1931.

Haken, H. 'Synenergetics', Berlin, Springer, 1977.

Honegger, B. 'Human paranormal phenomena as interhemispheric verbal communication: There is another system', PA Convention, 1979.

Hurst, S.L. 'Threshold logic', London, Mills & Boon, 1971.

Jordan, P. 'Verdrangung und Komplementaritat', Hamburg, 1947.

Kaufmann, S. 'Metabolic stability and epigenesis in randomly constructed genetic nets', Journal of Theoretical Biology, 22, 1969.

Kewney, G. 'Microprocessors don't always work', New Scientist, 18, 1978.

Klir, G. 'An approach to a general systems theory', New York, 1969.

Kornwachs, K. & Lucadou, W.v. 'Beitrag zum Begriff der Komplexitat', Grundlagenstudien aus Kybernetik und Geisteswissenschaft, 16, 1975.

Kornwachs, K. & Lucadou, W.v. 'Beschreibung und Entscheidbarkeit', Grundlagenstudien aus Kybernetik und Geisteswissenschaft, 17, 1976.

Kornwachs, K & Lucadou, W.v. 'Funktionelle Komplexitat und Lernprozesse', Grundlagenstudien aus Kybernetik und Geisteswissenschaft, 19, 1978.

Kornwachs, K. & Lucadou, W.v. 'Psychokinesis and the concept of complexity', Psychoenergetic Systems, 1978.

Lilly, J. 'Programming and metaprogramming in human biocomputers', Menlo Park, Whole Earth Catalog, 1967, 1970.

Lucadou, W.v. & Kornwachs, K. 'Can quantum theory explain paranormal phenomena?', in Morris, E. (ed) 'Research in parapsychology 1976', Metuchen, Scarecrow, 1977.

Lucadou, W.v. & Kornwachs, K. 'Fitting of PK results with a quantum mechanical model?', Parascience Conference, London, 1977.

Maddock, D.P. 'An electromagnetic method of enhancing the brain's coupling with telepathic psi', Parascience Proceedings 1973-1977, Part 2, Sprytown, Lifton, 1978.

Mattuck, R.D. 'Quantum mechanics, consciousness, and psychokinesis', Parascience Proceedings, 1973-1977, Part 2, Sprytown, Lifton, 1978.

Millar, B. 'The observational theories: a primer', EJP, 2, 1978.

Popp, F.A. 'Vom Wesen des Lebens', Umschau in Wissenschaft und Technik, 8, 1979.

Popper, K.R. & Eccles, J.C. 'The self and its brain', Berlin, Springer, 1977.

Pribram, K.H. 'A progress report on the scientific understanding of paranormal phenomena', in Shapin, B. & Coly, L. (eds) 'Brain, mind and parapsychology', Parapasychology Foundation, 1979.

Prigogine, I. 'Zeit, Struktur und Fluktuationen', Angew.Chem., 90, 1978.

Sarfatti, J. 'Implications of metaphysics for psychoenergetic systems', Psychoenergetic systems, 1, 1974.

Schmidt, H. 'A PK effect on prerecorded targets', J. ASPR, 70, 1976.

Schmidt, H. 'Can an effect precede its cause? A model of a noncausal world', Foundations of Physics, 8, 1978.

Walker, E.H. 'Foundations of paraphysical and parapsychological phenomena', in Oteri, L. (ed) 'Quantum physics and parapsychology', New York, 1975.

Walker, E.H. 'Comparison of some theoretical predictions of Schmidt's mathematical theory and Walker's quantum mechanical theory of psi', Journal of research in psi phenomena, 2, 1977.

Weiszacker, E.v. 'Erstmaligkeit und Bestatigung als Komponenten der pragmatischen Information', in Weiszacker, E.v. (ed) 'Offene Systeme I', Stuttgart, Klett, 1974.

Taylor, J.G. 'A review of scientific explanations for psi phenomena', Parascience Proceedings 1973-1977, Part 2, Sprytowm, Lifton, 1978.

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 3

NUMBER 4

MAY 1981

Robert L. Morris
Psychology Department
University of Edinburgh

EDITORIAL STAFF

Martin Johnson
Sybo A. Schouten

REGIONAL EDITORS

U.K.
Dr. John Beloff
University of Edinburgh
Department of Psychology
7 George Square
Edinburgh EH8 9JZ

Sweden
Dr. Rolf Ejvegaard
Mellansjovagen 35
S-14146 Huddinge

U.S.A.
Ms. Rhea White
2 Plane Tree Lane
Dix Hills
New York 11746

Australia
Dr. H.H.J. Keil
University of Tasmania
Box 252 C, G.P.O.
Hobart, Tasmania

W. Germany
Prof. Dr. Hans Bender
Prof. Dr. Johann Mischo
Institut fur Grenzgebiete der
Psychologie und Psychohygiene
78 Freiburg i. Br.
Eichhalde 12

Iceland
Dr. Erlendur Haraldsson
University of Iceland
Department of Psychology
Reykjavik

Italy
Dr. P. Cassoli
39, Via Valeriani
I-40134 Bologna

Prof. Dr. Emilio Servadio
Via di Villa Emiliani, 4-int.7
I-00197 Roma

The European Journal of Parapsychology publishes original experimental and theoretical papers on parapsychology. In addition, articles presenting reviews of literature, case studies and papers on related topics are accepted in so far as the subject can be considered relevant to the understanding of parapsychological phenomena or to the methodology of empirical research in parapsychology. Translations of papers originally published in a language other than English are also welcome.

Manuscripts should be written in English and should be typed, double-spaced throughout, with a wide margin, on one side of the page only. The organization of the paper must be clearly indicated by appropriate headings and subheadings. Include names of authors, positions and other affiliations, and mailing addresses. Each table

(continued inside back cover)

R. MORRIS
313 LINK HALL
SYRACUSE UNIV.

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

CONTENTS

Johnson, M.U.	Is the 'repeatable' experiment a sine qua non to establish parapsychology as a science?	p.315
Broughton, R.S. Millar, B. Johnson, M.U.	An investigation into the use of aversion therapy techniques for the operant control of PK production in humans	p.317
Bierman, D.J. Houtkooper, J.M.	The potential observer effect or the mystery of irreproduceability	p.345
Parker, A.	In defense of introverts: A critical note on supposed personality differences between ganzfeld experimenters	p.373
Pasricha, S.K. Barker, D.R.	A case of the reincarnation type in India. The case of Rakesh Gaur	p.381
Ejvegaard, R.	Parapsychology and repeatability	p.409
Sargent, C.L.	The repeatability of significance and the significance of repeatability	p.423

PUBLICATION POLICY

Twice a year the Parapsychology Laboratory of the University of Utrecht publishes the European Journal of Parapsychology. The object of the European Journal of Parapsychology is to stimulate and enhance the activity in this field, especially in our corner of the world, by communicating research results and issues related to professional parapsychology. Although there will be an emphasis on experimental work, theoretical articles are also welcome. Contributions from all over the world will appear in the journal.

A hallmark of the European Journal of Parapsychology is the attempt to avoid selective reporting, that is, the tendency to bury 'negative' results and only to publish studies that 'turn out'. To avoid turning the journal into a graveyard for all 'unsuccessful' studies, we require that the acceptance or rejection of a manuscript should take place prior to the phase when the experimental data are collected. The quality of the design and methodology and the rationale of the study are considered more important than the level of significance of the outcome of the study. As a practical rule, we advise the potential contributor of an article to submit the design of his planned study before the work is actually carried out. The rationale of the study should be stated, as well as all the hypotheses related to it. Furthermore one should try to specify the number of subjects, the number of trials, etc., plus the type of statistical methods one plans to use for evaluation.

Priority will be given to the publication of studies which fulfil the above-stated publication policy.

The final manuscript with presentation of results must reach us two months in advance of the official publication dates, which are May 1st and November 1st.

IS THE 'REPEATABLE' EXPERIMENT A SINE QUA NON TO ESTABLISH PARAPSYCHOLOGY AS A SCIENCE?

Over the years there has been a great deal of debate over the existence of something approximating a 'repeatable' experiment in parapsychology. Most knowledgable persons would probably agree that there isn't yet a method that can assure you of demonstrating psi whenever you like.

The reason why the repeatability-issue has attracted much interest and attention among parapsychologists as well as among the critics is undeniably the fact that if a very high degree of repeatability were at hand, especially if there were a rather high degree of inter-laboratory repeatability, one could claim that a demonstrable experiment had been attained. If one really had attained a demonstrable experiment, the case for the existence of psi would be a more or less clear-cut case.

In addition, there is at least one more quite important aspect of having a truly repeatable experiment: the main methodological importance of repeatability is the possibility for improvement of measurements. Such improvements are instrumental for theory-testing and theory-building. It should be remembered that repeating an experiment just for the sake of repeating it is scientifically as vacuus as the child's request that the 'magician does the miracle again'. A special case of reproduceability of psi-results would be to find that there are no extra-chance variations at all in a long series of studies. Only results fully explainable in the light of chance-expectance were obtained.

Rather independently of one's assessment of whether or not there is

a 'sufficient' degree of repeatability in parapsychology, the very concept of 'repeatability' is a quite interesting one from a strictly epistemological point of view. Because of the importance of this issue we decided to invite potential authors to submit papers on the concept of repeatability in science for publication in the E.J.P.. From the beginning we had hoped to bring about a special issue of the E.J.P., entirely dedicated to the problem of repeatability. It turned out, however, that several scientists and philosophers of science who have written on repeatability or criticized parapsychology for not having repeatable experiments, did not feel that they had much to contribute. (One leading expert on methodology within the behavioural sciences flatly admitted that "honestly speaking I do not think I have much to add to what previously have been said".)

In this issue we are publishing two of the contributions we have received. One is by Dr. Carl L. Sargent, an adament spokesman for the view that we have acquired a reasonably high level of repeatability in at least a few types of parapsychological experiments. He certainly seems to adhere to Feyerabend's general view of what science is about. The other paper is written by Dr. Rolf Ejvegaard who follows a rather original line of thought in relation to the concept of repeatability in science. As one can see, the problem of how to distinguish between what is science and what is pseudo-science is not a simple task.

It is to be hoped that the readers will be stimulated sufficiently to ponder upon this rather central concept in experimental science, and that they also will share their ideas by either submitting comments on these articles or full papers to be published in the next issue of the E.J.P.. The deadline for such a contribution is September 15th of 1981.

Martin Johnson

AN INVESTIGATION INTO THE USE OF AVERSION THERAPY TECHNIQUES FOR THE OPERANT CONTROL OF PK PRODUCTION IN HUMANS

Richard S. Broughton
Brian Millar
Martin Johnson
University of Utrecht

It has always been an assumption of parapsychological research that the demonstration of psi is subject to various constraints and controls within the individual: It is to the specification of these constraints and controls that research is largely directed. This is, of course, a natural position for the field, closely modelled as it is on psychology, and the assumption that paranormal behaviour parallels other forms of human behaviour is quite justifiable given the traditional interpretation of parapsychological evidence.

How successful have parapsychologists been in identifying those factors which are responsible for governing the production of paranormal behaviour? An impressive body of literature claiming to document various lawful relationships between psi and variables known to govern other forms of human behaviour has grown up over the years. However, given the continuing cries of critics for evidence of a more tangible sort than has so far been forthcoming (e.g. Wheeler, 1979) and the pressing need for the reasonably stable source of psi required to test some of the more promising recent theories (see Millar, 1978)

Note: This paper fulfils the publication policy of this journal.

it is not unreasonable to apply a more utilitarian criterion. Just how does one go about getting psi to appear in the laboratory?

To the number of parapsychologists who seem to have little difficulty in finding evidence of psi in their work this question may appear almost facetious. To the many parapsychologists who rarely, if ever, see any evidence of psi this question, far from seeming facetious, stands out as one of the foremost problems. The continuing inability to specify the requirements for even modestly reliable production of psi phenomena under controlled conditions represents a strong indictment of the field's alleged progress.

Recently there has been considerable optimism expressed by some parapsychologists in the area of psi and internal noise reduction. In particular, the ganzfeld technique (Honorton, 1977; Sargent, 1979) is, according to those workers, well on the way to providing the specifications for the reliable production of psi in the laboratory. Pratt (1978), however, cogently reminds us of the final oblivion of so many other such "most promising" lines of research in the past, and we must bear his admonition in mind, even while cherishing the hope that the ganzfeld technique may prove to be an exception and thus completely counter our foregoing criticisms.

In recent years a new interpretation has been competing for the attention of those who seek to make sense of parapsychological findings. Depending on one's point of view it is either the intractable problem or the exciting discovery of psi based experimenter effect. This would have most of the seemingly successful lines of parapsychological research with unselected subjects interpreted not as more or less repeatable experiments but as the products of more or less repeatable experimenters. The evidence for this interpretation (Kennedy and Taddonio, 1976; White, 1976a,b; Broughton, 1977; O'Brien, 1979) cannot be dismissed lightly and further Millar (1979) one of the strongest proponents of this view, has argued that in the absence of psi-producing experimenters there is very little evidence of psi in the general population. If parapsychologists were regularly to find evidence of psi clearly and reliably linked to identifiable activities on the part of the subject then the question of interpretation, 'whose psi?' would not arise. Such evidence would, of course, have to be of the sort which could be gathered by various investigators and not linked to a particularly complex or idiosyncratic manner of conducting research that would provide forever plausible excuses for replication failures.

An especially tantalizing candidate for such a source of reliable psi is the possibility that paranormal behaviour can be learned, developed, or in some way brought under the control of the individual. With regard to one type of psi ability, ESP, the idea that this could be learned has received intermittent consideration throughout the history of the field. However the finding that, if anything, most subjects seemed to decline rather than improve in their performance tended to discourage such speculations. An important paper by Tart (1966) pointed out several fundamental problems regarding the learning of ESP in forced-choice situations and in 1975 he published the results of a major attempt to overcome these problems.

Although an ambitious undertaking it remains an open question whether the Tart (1975) study did in fact demonstrate the learning of ESP as the author claims. Stanford (1977b) has catalogued a large number of theoretical and methodological weaknesses and a more recent paper by Gatlin (1979) has raised the question of to what extent the subject may take advantage of short-term non-randomness in the target sequences in ESP experiments using immediate feedback. A further large scale study by Tart and his collaborators (Tart, Palmer and Redington, 1979) added little to the earlier work.

Attempts to demonstrate the learning of PK would largely bypass many of the problems inherent in the use of ESP but unfortunately very little work has been done in this area. Although the requirement of immediate feedback has been met in many studies and quite a number of these have used talent-selected subjects, Stanford (1977a) in his encyclopaedic review of PK was unable to adduce any evidence of learning effects in the research to date. We are aware of only two studies which have specifically addressed the question. In the first (Thouless, 1945) a decline was found. Although an incline suggestive learning appeared in one of the many analyses in the first experiment of the other piece of work (Camstra, 1973) no such effect was seen in the follow-up.

Germane to any discussions of learning paranormal behaviour is the question of motivation to learn. There is no evidence suggesting that whatever paranormal abilities may exist are easy to learn. Rather the opposite is suggested. However, rarely in experimental situations is the subject's motivation any higher than to gain the experimenter's approval or generally to "do well" in the task. It is quite possible that a subject may have to have very good reason indeed to learn, or even to use consistently, whatever paranormal abilities he or she may

possess.

The Behaviourist approach to psychology has developed one particular technique of bringing about learning which has been repeatedly demonstrated to be a highly reliable means of getting subjects to emit specific behaviours. The technique is operant conditioning and in its therapeutic form, as behaviour modification, it has been shown to be successful in changing or eliminating particularly entrenched human behaviours which are deemed maladaptive. Although it must be admitted that this area of research is particularly complex and there is considerable debate among its practitioners concerning the details of the learning which takes place under conditioning paradigms it must also be pointed out that operant conditioning in general and its behaviour modification form in particular have had records for getting results which are enviable in psychology, not to mention parapsychology.

It is not the intention of this report to review the successes and problems involved in human behaviour conditioning. Among many reviews of such material one may consult Beech (1969), Kanfer and Phillips (1970), or for a condensed summary Hilgard and Bower (1975). The present investigation, which developed from a long-standing idea on the part of one of the authors (M.J.), had as its purpose to see if one particular form of human conditioning, aversion therapy or punishment, can enable subjects to gain control of or learn to use PK.

Although many early researchers, including Skinner himself, were skeptical of the efficacy of punishment in changing or eliminating behaviours, a number of reviews in the mid 1960's (Solomon, 1964; Azrin and Holz, 1966 among others) have largely changed the climate of opinion regarding punishment. Currently it is in wide use in behaviour therapy programs. Extensive research has clarified the various parameters for effective utilization of punishment and Azrin and Holz (1966) have catalogued the requirements for those who wish to employ this technique.

Among the most effective punishments contingent upon an undesired response is a painful electric shock. The behaviour therapy research literature abounds with examples of the use of electric shock in the successful treatment of a large range of behaviours, ranging from sexual disorders, alcoholism and obesity to certain aspects of autistic behaviour and aggression in retardates. It is clear that a painful electric shock is something which subjects will go to great

lengths to avoid.

The purpose of this investigation was, quite simply, to see if a painful electric shock contingent upon a predefined "unacceptable" level of performance in a PK task is sufficient to cause subjects to avoid such responses to a degree that would indicate PK activity. In other words, can PK be conditioned by one of the best conditioners of behaviour currently available?

For the purposes of this initial investigation it was necessary to adopt a "black box" approach with regard to the precise nature of the "learning" which might be demonstrated. The pragmatic approach of early behaviourism governed this investigation and the goal was to see if punishment by electric shock could reliably eliminate the occurrence of low PK scores. These would, of course, have to be achieved through paranormal means. As will be seen below the "behaviour" which the subject must learn to avoid is the chance occurrence of low scores. The nature of the PK task provides the subject with ample indication of whether or not a particular run is dangerously close to shock level. Obviously various types of learning could be involved as well as different PK strategies. For example the learning could involve early discrimination between chance "safe" runs and those which require a PK boost. Another possibility is that the subject exerts a PK effect on all trials to be on the safe side.

It is clear that the present investigation confounds learning with the possible generalized effect on PK performance of a particularly powerful motivator, hence the "black box" approach and the hope that this initial investigation might provide the data to guide future efforts.

While it was the hope of the investigators that this extreme form of conditioning would be successful in securing reliable psi effects it was also noted that a negative result in this experiment would not be without utility. The inability to condition paranormal behaviour with one of the most effective conditioners available would at least raise the question of whether perhaps psi functions in a manner entirely alien to other human abilities and would suggest that investigative approaches based upon the traditional assumptions of psychology might not be appropriate.

METHOD

The usual approach in the testing of conditioning effects in behaviour therapy is to use an ABA design with a small number of subjects or patients. The A condition represents the baseline or free operant measure of the dependent variable. In the B condition the experimental manipulation is introduced and the dependent variable is again measured. The terminal A condition is a return to the pre-manipulation or baseline condition with a third measure of the dependent variable. Of course it must be agreed with Paul (1969) that such designs should serve only as adjuncts to factorially designed studies with large groups but it is clear that in this case researchers will need convincing evidence of the efficacy of the technique before studies of the latter type are likely to be carried out.

Thus this investigation was conducted according to what amounted to a case-study approach. There were several reasons, apart from practical considerations (such as getting subjects), which guided this choice. The first consideration was the obvious one that not all persons may have PK potential in the first place and for these no amount of conditioning will be successful. A second consideration was that, in therapeutic applications of behaviour modification, for some persons the technique just does not work. A third reason was that even if successful avoidance of shocks occurred in several cases the actual manner of achieving this end might be quite individualistic and each case would merit individual examination.

The general method used in this investigation was as follows: Subjects performed PK runs on a computer-controlled PK game device called Kop van Jut which had been used in previous research. This game mimics the test of strength found in carnivals and amusement parks and consists of a number of lights in a column, topped with a bell, which are illuminated successively from the bottom for PK hits. (For a detailed description of this device see Broughton (1979).) There was a predefined point on the Kop van Jut scale below which any run score would be considered an unacceptable response. In the A condition scores falling below the criterion had no aversive consequences but in the B condition each and every below criterion run score was punished by a shock administered according to the guidelines of Azrin and Holz (1966).

Each subject was given the ABA treatment and after a pause of about one week each subject repeated the ABA treatment as a replication. A replication was required for two reasons. Owing to the nature of the work robust, repeatable effects were wanted or else there would be little point in using such an extreme method. Secondly, if highly individualistic PK strategies emerged from the data it would be necessary to confirm them in separate data before they could be considered real. It was planned that both original and replication sessions were to be completed before any data were examined. However, for practical reasons noted below, this plan was not fully followed. It should be obvious that the replication could not have been made contingent upon the success of the original series because it would then have been in the subject's long term interest (in avoiding shocks) to allow the original to fail.

EQUIPMENT

The PK game was "played" on the above mentioned Kop van Jut apparatus which is described elsewhere. Through an interpreting device the Kop van Jut apparatus was controlled by a computer program running on a PDP-11 computer. The general appearance of a run on this device was that, upon the press of a button, a column of lights began lighting from the bottom upwards, very rapidly at first but with decreasing speed until it came to a stop at some point, hesitated for a moment, and then fell back. The display was meant to simulate the trajectory of a weight rising and falling. Alongside the column of lights was a strip of paper bearing various legends such as "Average", "Not Bad", etc. Below lamp number 22 the paper was colored red indicating the shock zone and above that it was green. A sketch of the device may be seen in Figure 1.

Random binary numbers for the targets were generated by a noise-diode based random event generator coupled to the computer. Control tests of REG functioning were built into the program such that each time an experimental target was obtained from the REG the computer also obtained and stored a control target which was not displayed. Thus each target which governed the display to the subject was matched, one for one, with a control target obtained at virtually the same time. In this experiment (in contrast to Broughton (1979)) all targets were generated in real-time: none were pre-recorded.

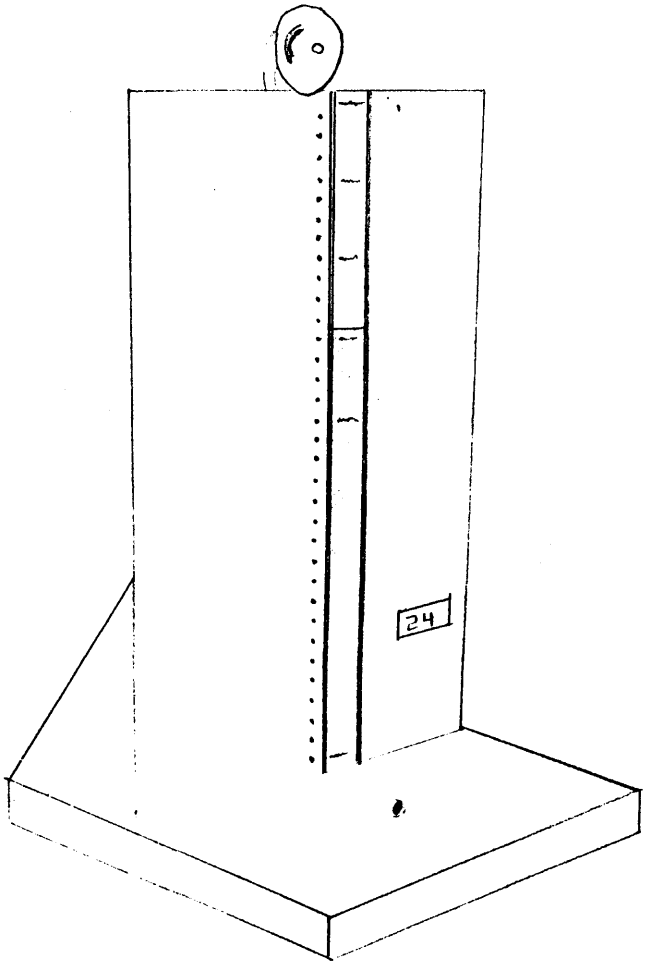


Fig. 1 - The "Kop van Jut" device

Shocks were administered automatically by the computer via a commercially manufactured shock device (Model DSU, Tau Electronics, Ltd., Edinburgh) used in behaviour therapy applications. Shocks were administered to the back of the wrist through wrist-band electrodes.

EXPERIMENTAL CONDITIONS

Each button press by the subject initiated a run of 50 binary trials. A session consisted of 24 such runs.

The A or baseline condition consisted of eight sessions on the Kop van Jut apparatus. No punishment was administered for below criterion runs. Subjects were required simply to play the game. Previous research (Broughton, 1979) had demonstrated that this game, while generally enjoyable for the subjects, is not of itself effective in enhancing PK scoring.

For the B or experimental condition subjects did another eight sessions but in this condition all run scores falling below 22 were punished with a shock. At this level, by chance alone, the subject could expect about four shocks per session, i.e. with a probability of about 1/6 (0.1611 from binomial expansion) per run. This level was set mainly on the basis of the intuitive feeling that more than four per session would discourage participation and at the same time it was low enough for the optimistic subject to feel that for any particular session he might escape shocks entirely.

The second A or return to baseline condition required subjects to perform the task without the threat of shocks.

No other requirements were made of the subjects. Within reason they were free to do sessions whenever they felt up to it. For both the A and B conditions every effort was made to accommodate the time for running to the subject's wishes.

SUBJECTS

The authors of this paper served as subjects along with one member of the staff of the parapsychology laboratory. All subjects were

thoroughly aware of the aims of the experiment as well as the procedure to be used. All subjects signed a statement of informed consent. Originally it was hoped to have between six and eight participants but in the end it proved possible to enlist only four.

It should be recorded at the outset that at least two of the subjects (R.B. and M.J.) have on occasion demonstrated effects which appear to have been paranormal in nature so it would not seem plausible to argue, in the event of a negative outcome, that the subject pool was entirely devoid of paranormal potential.

PROCEDURE

As all subjects were well acquainted with the aims of the experiment no formal introduction was necessary. Likewise all subjects were familiar with the operation of Kop van Jut.

The running of a subject's series of sessions was arranged to suit his convenience, the only restriction being that ideally not more than one session should be done in any three-hour period. The preferred rate was one session per day. Each subject maintained a record of visits although no record of performance other than that stored in the computer was maintained.

A typical session started with the subject initiating the controlling program for the apparatus. He seated himself comfortably in the small room with Kop van Jut and when ready played the game ad lib. A counter on the device informed the subject when the session was complete, at which point he was free to leave. Data were automatically stored upon completion of a session.

Before each shock session the subject sampled the shock to insure that the intensity level was correctly set. (The intensity was as high as the subject could stand.)

The B sessions were conducted in a manner similar to the A sessions except that prior to beginning the subject was instructed to wash his hands, particularly the wrists, and dry them throroughly. Before starting the session the wrist electrodes were attached and connected to the shocker. The hand shocked was generally alternated through the series.

At the conclusion of one ABA series there was a rest period, usually of about one week, before the replication was embarked upon. The replication was identical to the initial series. No data were to be examined until all subjects had completed both series. As noted below, however, this did not prove practicable.

ANALYSIS

The principal dependent measure to test the effectiveness of the treatment was the number of run-scores below 22. In addition, general scoring effects, variance effects and possible release-of-effort effects were of interest.

Novel use was made in this experiment of one particular analysis, which may prove of value elsewhere. It would be naive to assume that when PK occurs that it manifests only in accordance with the conscious goal of the subject: on the contrary a diffuse scattering of secondary effects may be expected. Consequently some global measure of patterning in the otherwise random sequence of targets might detect PK missed by a test for target biasing only. For this purpose, Good's serial test, frequently used in the context of checking the randomness of random number generators, was employed here as a detector of PK-generated patterning.

Because of the blind pilot - confirmation collection of data used in this experiment the investigators applied a range of tests in the hope of covering all reasonable possibilities for psi effects. In order to avoid the problem of a statistical scrounging artifact no effect could be considered demonstrated unless it was replicated in the data of the confirmation series. The analysis of the second set was not undertaken until all pilot analyses were deemed complete.

The reason for the choice of such a stringent criterion for the demonstration of effects as well as the decision to employ a conventional replication at a later time instead of the easier "Edinburgh Split" method stemmed from the investigators' belief that the rather extreme methods used in this project require a particularly robust success if they are to be recommended to other investigators.

PILOT STUDY

Due to the deadline for the preliminary submission of data for the 1979 Parapsychological Association Convention, at which a preliminary report of this research was presented, it was not possible to meet the desired requirement that all data be in before any analysis. Data analysis of only the pilot sessions commenced while the participants were at various stages of the confirmation series.

One of the analyses was to examine for the possibility of a release-of-effort effect. The r-o-e period was defined as all trials taking place after the 22nd lamp was lit. Before that point it could be assumed that the subject was exerting an effort to avoid a shock. After the 22nd lamp was lit the subject knew he was "safe".

As a reminder of the structure of the experiment it should be noted that there were three conditions: Baseline, shock, and return to baseline. Each condition consisted of eight separate sessions with the PK apparatus. Each session consisted of 24 runs of 50 binary trials. Thus for each conditon there were 192 runs for a total of 9600 binary trials.

RESULTS

The results of the four subjects will be treated individually.

For subject R.B. the number of run scores falling below 22 for the baseline condition was 28, for the shock condition, 30, and for the return to baseline condition, 32. The chance expectation for this measure is approximately 31. Thus subject R.B. received 90 shocks in toto and none of the conditions differed significantly from chance.

No evidence of significant above chance scoring was found in the data of the three conditions with the baseline producing 4812 hits, shock, 4798 hits, and return to baseline, 4735 hits, none of which differed significantly from the mean chance expectation of 4800.

The release-of-effort rate of scoring did not differ significantly from the effort rate in any of the three conditions. Variance measures also were well within chance limits. The pilot results for subject

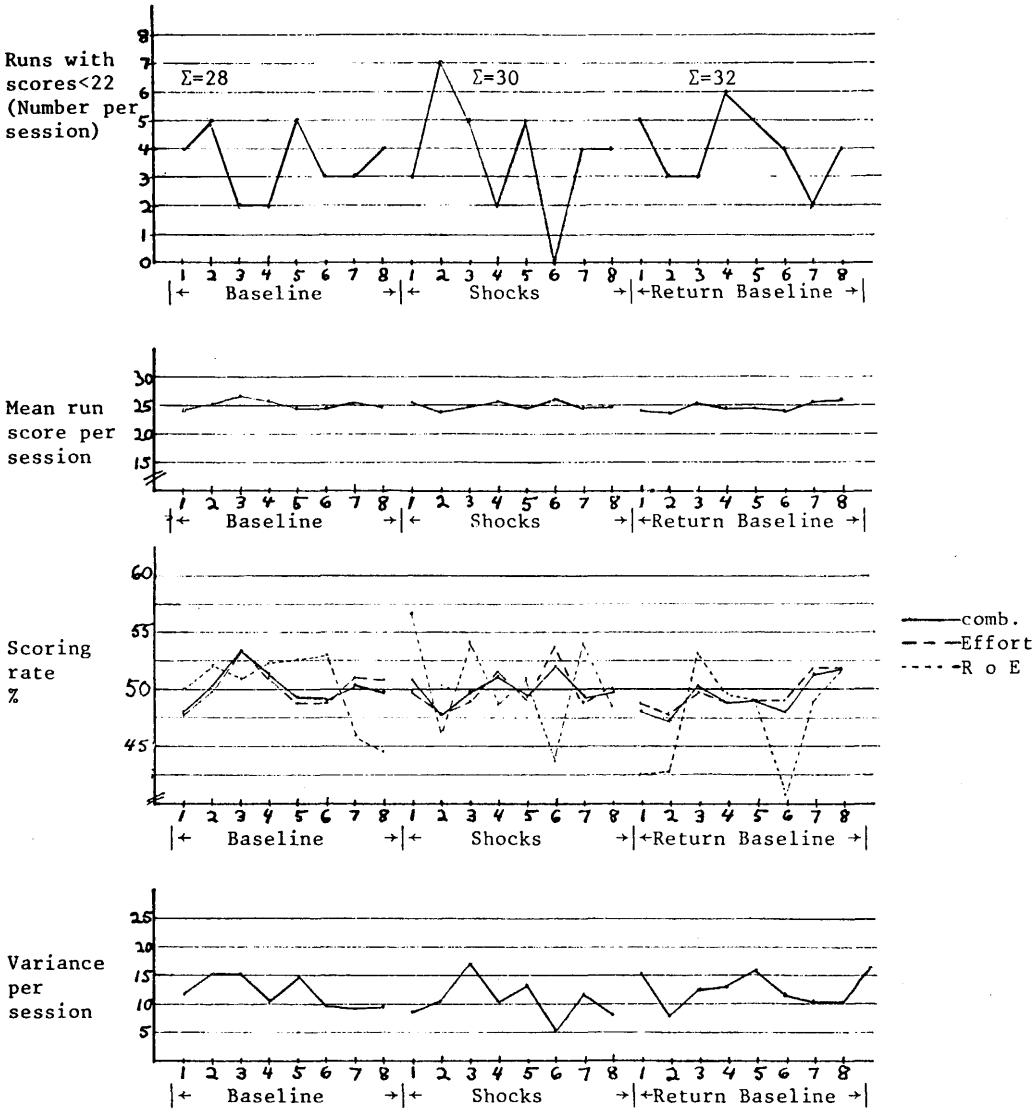


Fig. 2 - Pilot results for subject R.B.

R.B. are presented graphically in Figure 2.

For subject B.M. the number of run scores falling below 22 were, in the order baseline, shock, return to baseline, 43, 24, 32. The shock condition result of 24 shocks is below the expected number but not significantly so. The first baseline result of 43 "shocks" is significantly higher than chance expectancy with a two-tailed binomial probability of $p<.02$. The shock condition produced significantly fewer below-22 scores than did the preceding baseline, CR(diff)=2.57, p=.01 (2-tailed).

The first baseline condition, in which the high number of below-22 scores were noted, produced a significant negative deviation in the overall scoring with 4703 hits for a CR=-1.97, $p<.05$ (2-tailed). In this condition the effort score was 4128 hits (MCE=4222) for a CR=-2.04, $p<.05$ whereas the r-o-e score was 575 hits (MCE=578) for a CR=-0.15). Neither of the other two conditions produced significant departures from chance with the shock condition yielding 4787 hits and the return to baseline condition yielding 4838. Variance measures for all conditions were within normal limits. The data for subject B.M. are presented graphically in Figure 3.

For subject M.J. the number of run scores falling below 22 were all within chance limits with totals of 26, 34, 26 for the baseline, shock and return to baseline conditions respectively. Overall scoring results were within chance limits as well, with totals of 4836, 4768, and 4765 for the three conditions.

In the scores for the shock session there were data bearing on the release of effort hypothesis. This subject provided the only interesting data regarding the r-o-e effect in that during the shock sessions the effort trials yielded Hits=4144 (MCE=4207) for a CR=-1.36 while the r-o-e trials produced Hits=624 (MCE=593) for a suggestive CR=1.77, p=.077 (two tailed). Using the scoring rates for the two types of trials for each of the eight sessions a Wilcoxon test failed to find a significant difference between the effort and the r-o-e trials. It is also worth noting that before the third shock session the subject reported that the pre-session sample shock was particularly strong but he would leave it at that level. This observation was entered into the record before the runs started. For that session the effort trials produced a CR=-0.67 but the r-o-e trials produced a CR=2.66, p=.008 (two tailed).

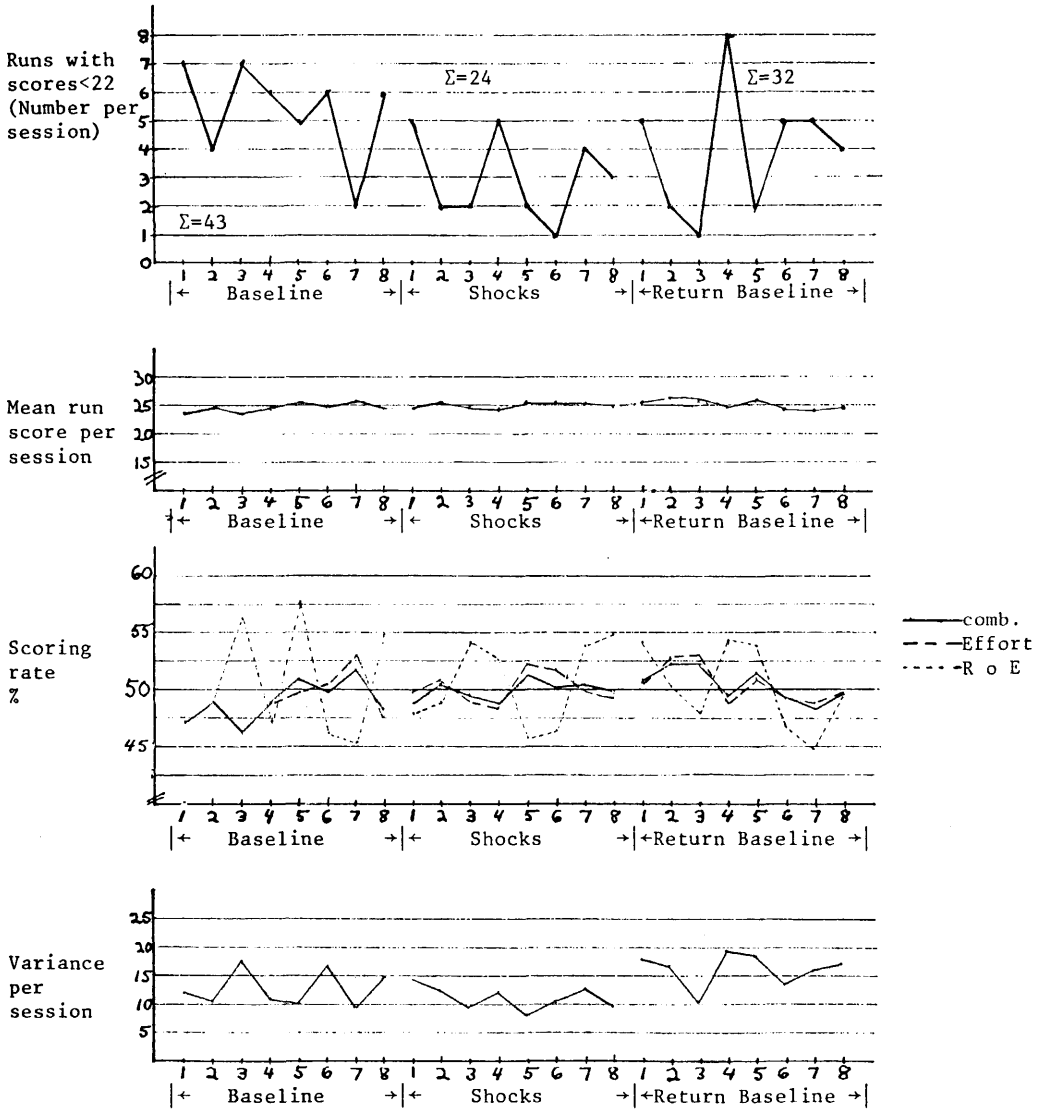


Fig. 3 - Pilot results for subject B.M.

Variance data in all conditions were within normal limits. The results for subject M.J. are presented graphically in Figure 4.

For subject S.S. the number of run scores falling below 22 were within chance limits with the baseline yielding 26, shock, 31, and return to baseline, 32. Overall scoring results similarly were within chance limits. Hits for the baseline, shock, and return to baseline conditions were 4820, 4801, and 4838 respectively. Release-of-effort rates of scoring did not differ significantly from effort rates in any of the conditions nor did the variance fall outside normal limits. The results for subject S.S. are presented graphically in Figure 5.

As a final test for any extra chance patterning of the data which might indicate a PK effect, randomness tests were performed on the binary data from each session in all conditions. Good's Generalized Serial Test was used to examine for second, third and fourth order effects. Although occasional modest deviations from chance were noted in the Chi-square results the frequency of these, both for the experimental and the matched control data, was well within expected limits considering the number of individual tests performed.

For brevity the results of the matched control data as well as that from routine tests of RNG functioning are not reported in detail here. In all cases these data indicated that the RNG was functioning adequately in the absence of subject efforts.

INTERIM

From the pilot study results it was clear that this experiment failed to achieve its desired objective of causing subjects to reduce by PK the number of shocks administered. The nearest this goal was approached was in the data of subject B.M. where fewer shocks were received than would be expected by chance but the total fell short of significance. Subjects R.B. and S.S. received almost exactly the chance level of shocks and subject M.J. received slightly more.

Although it was demonstrated that subject B.M. received significantly fewer shocks in the shock condition than in the preceding baseline condition this was largely due to the subject receiving a significantly large number of "shocks" in the baseline condition. This does not appear to be the sort of result which could

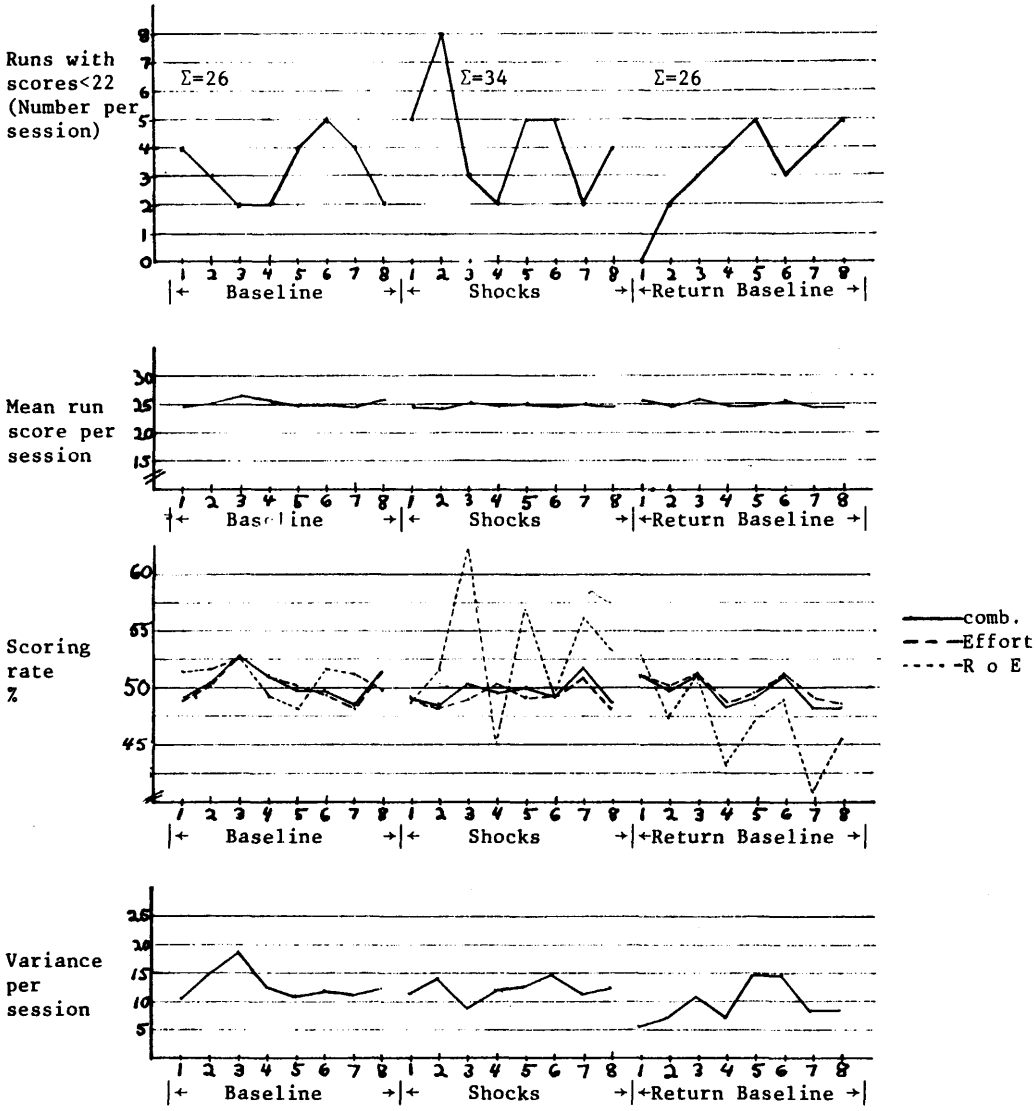


Fig. 4 - Pilot results for subject M.J.

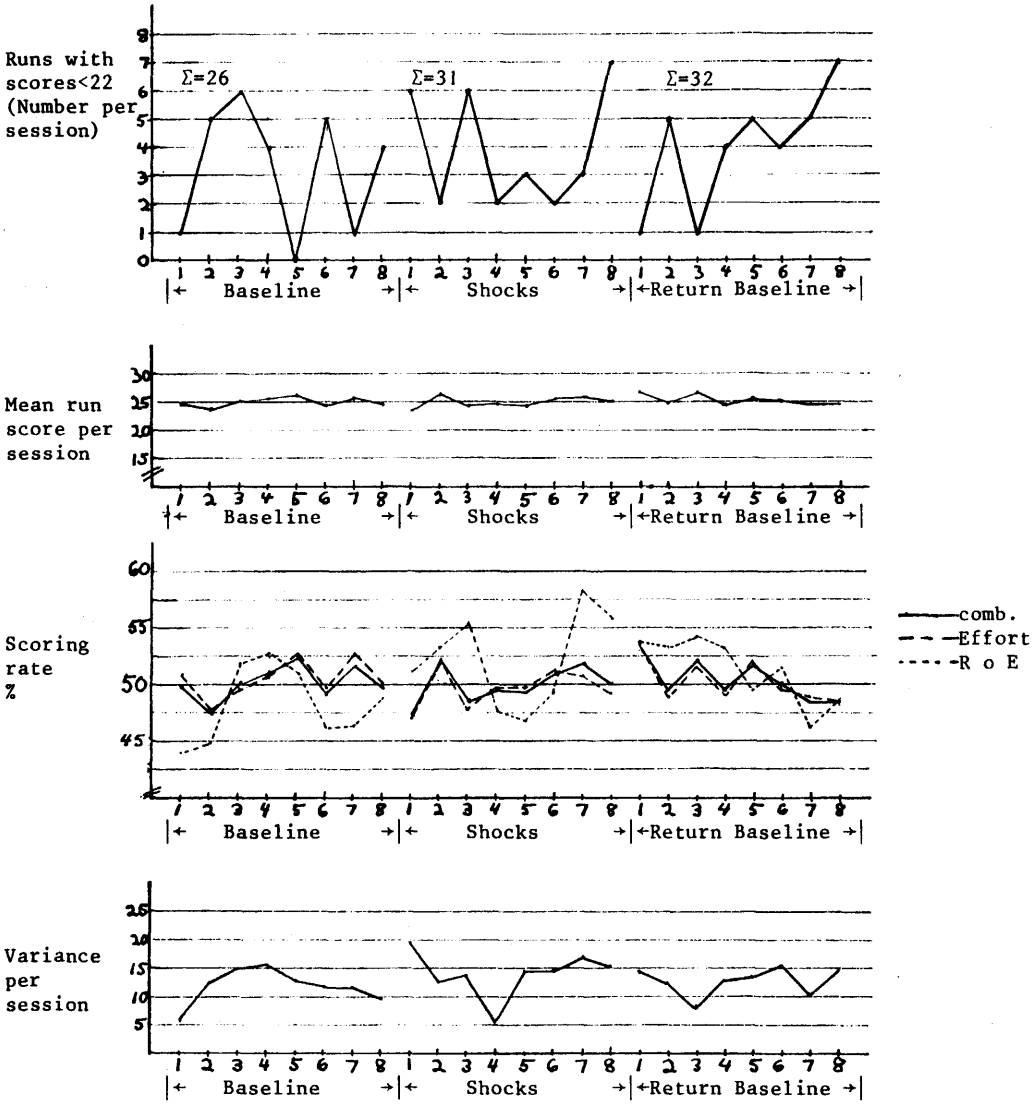


Fig. 5 - Pilot results for subject S.S.

be considered supportive of the hypothesis that shocks will condition PK learning but it seems possible this could be a genuine PK effect and it was tabled for examination in the confirmatory data.

The data of three of the subjects provided nothing to suggest a release of effort effect despite the fact that the experiment provided a very clear dichotomy between effort and release of effort. The data from subject M.J. did suggest that at least for him there may have been a release of effort effect present while under the threat of a shock. This impression was strengthened by the one session where the subject found the shock so strong that he felt it necessary to record the fact prior to starting and which in the end produced a result indicating a much higher scoring rate for the r-o-e trials. It was of considerable interest to see if this effect would be found in the shock sessions of the confirmation study.

CONFIRMATION STUDY

RESULTS

Subject S.S. was delayed in completing the confirmation study and was overtaken by a major change in the computer hardware during his shock condition. Consequently data from only three subjects are available for the confirmation. This is no great loss since for S.S. no suggestive features had been found in the pilot data, so no predictions had been made for the confirmation.

The pilot for subject R.B. also had given no evidence of any PK effects so none were expected in the confirmation: none were found. The numbers of runs with scores below 22 were 30, 28, and 28 for baseline, shock, and return-to-baseline respectively. No evidence of above chance scoring in any of the conditions was found. The results for subject R.B. are presented graphically in Figure 6.

For subject B.M. the pilot results had indicated a significantly high number of below 22 runs in the baseline condition. This was not replicated in the confirmation data. His numbers of runs below 22 were 25, 28, and 34 for the three conditions and the baseline score was the lowest of the series rather than the highest as in the pilot. The number of shocks was close to chance as were all other measures of

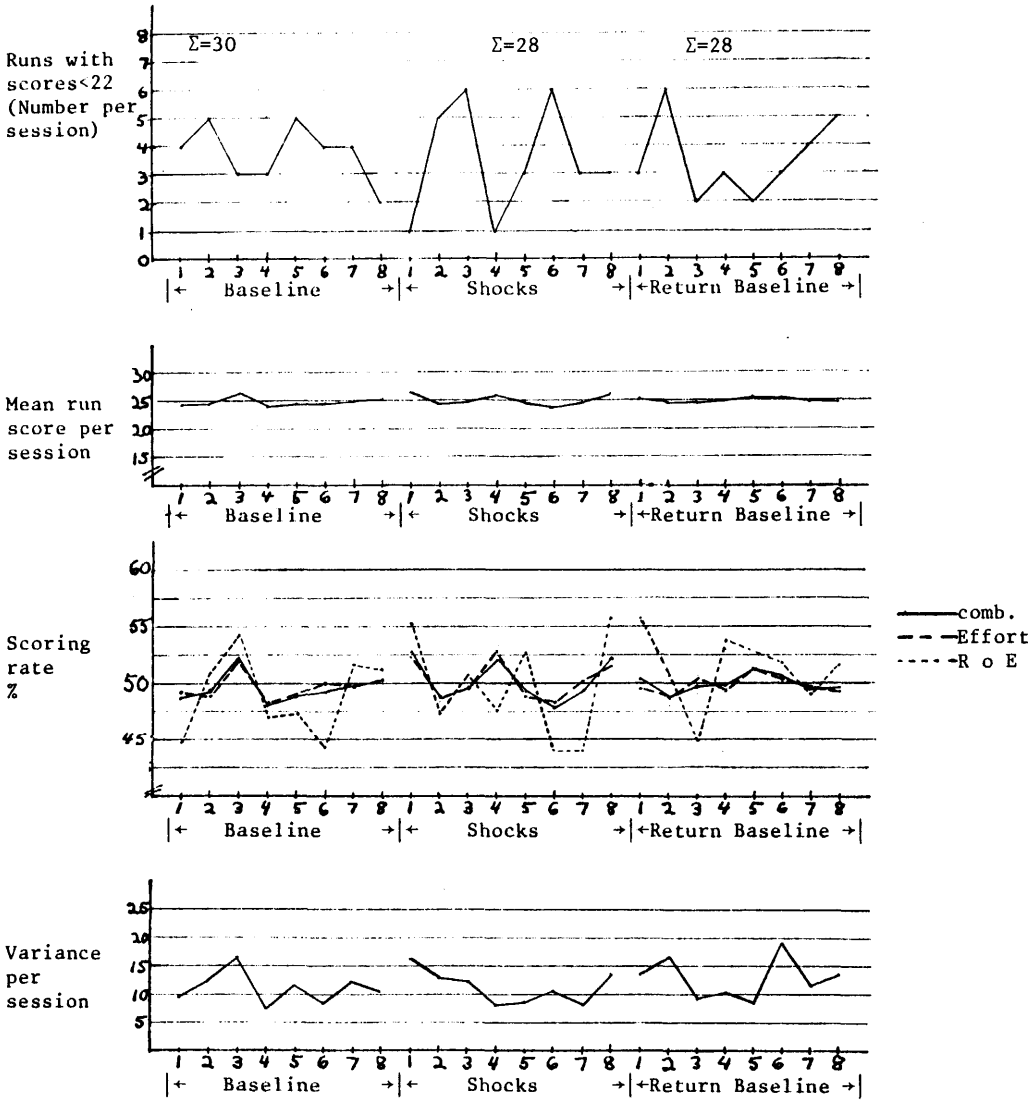


Fig. 6 - Confirmation results for subject R.B.

scoring. These data are presented in figure 7. Because of the premature analysis of the pilot data, B.M. had necessarily become aware of the pilot findings prior to the confirmatory study. For what it may be worth he reports that this knowledge did not noticeably change his psychological approach to the task.

For subject M.J. the number of below 22 scores were all close to chance with 33, 32, and 29. Unfortunately the overall scoring rate for the shock session gave none of the suggestion of a release-of-effort effect that the pilot data did. Effort trials yielded a CR=0.81 and the r-o-e trials gave CR=-0.34. In the confirmatory shock series the opportunity arose to introduce a manipulation similar to the 'session 3' effect noted in the pilot. (The subject was completely unaware of the pilot results at the time.) After one session the subject reported that the shocks seemed weak and he thought that the battery should be examined. Before the next session this was done and the shock intensity was increased considerably. The subject then reported that the shock was very strong. An examination of subject M.J.'s data in figure 8 might suggest that this session was number 2 where the r-o-e trials yielded a CR=2.30 ($p=0.02$, two tailed) in contrast to the effort trials which gave CR=-1.00. This is incorrect, however, since the shock intensity was increased just prior to session 5 which provided the lowest r-o-e score of the series with CR=-1.39. All other tests for above chance scoring yielded chance results.

Since the application of Good's Generalized Serial Test to the pilot data yielded no effects for prediction in the confirmation data this test was not run. As in the pilot study the matched control data for all subjects indicated satisfactory performance of the RNG in the absence of subject efforts.

DISCUSSION

In this experiment the pilot study provided, at best, only tenuous evidence suggesting the presence of psi and virtually no indications of conditioning or learning effects for psi ability. The few suggestions of psi in the pilot data failed to be confirmed in the confirmation part of the experiment.

The main conclusion of the experiment must, therefore, be a negative one. Aversion therapy techniques, in this case punishment for low

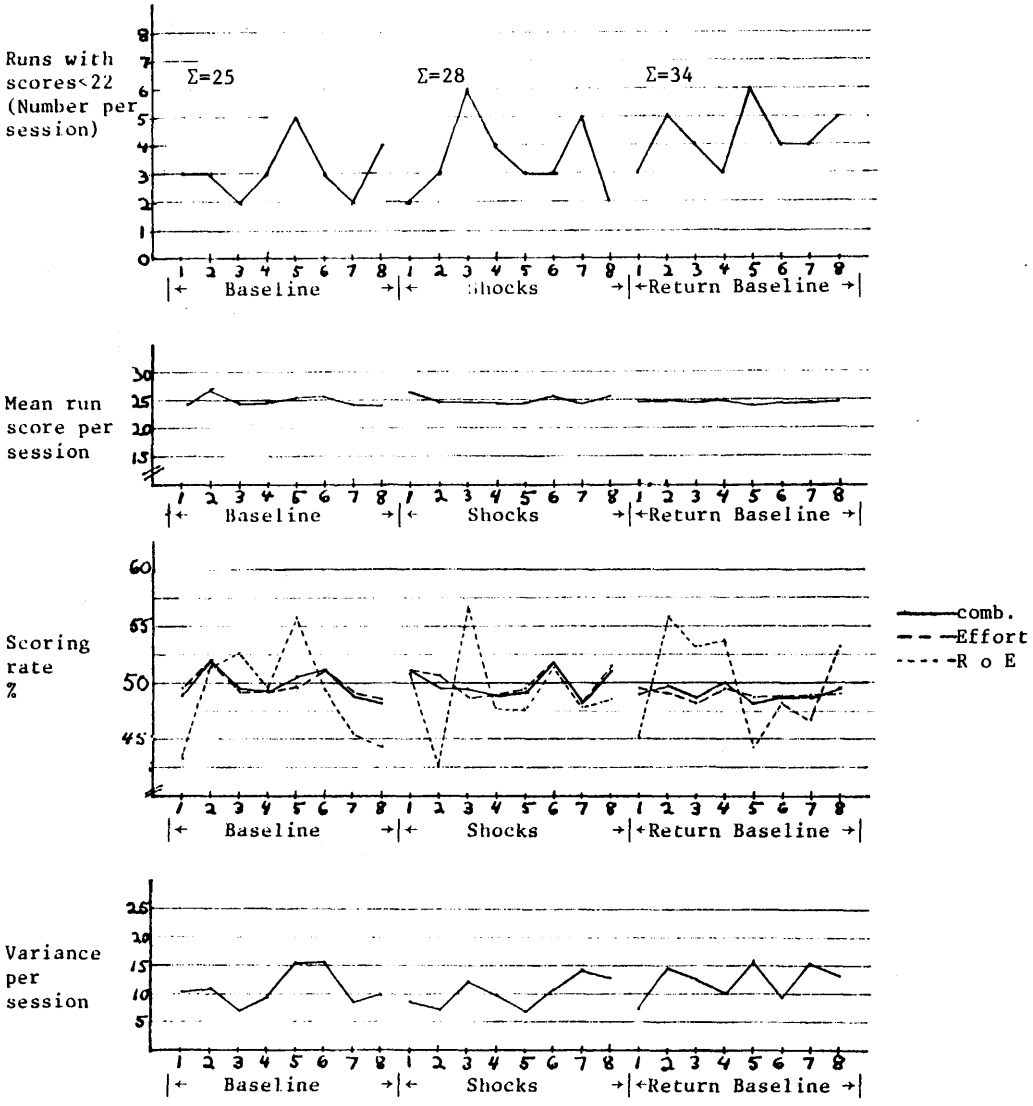


Fig. 7 - Confirmation results for subject B.M.

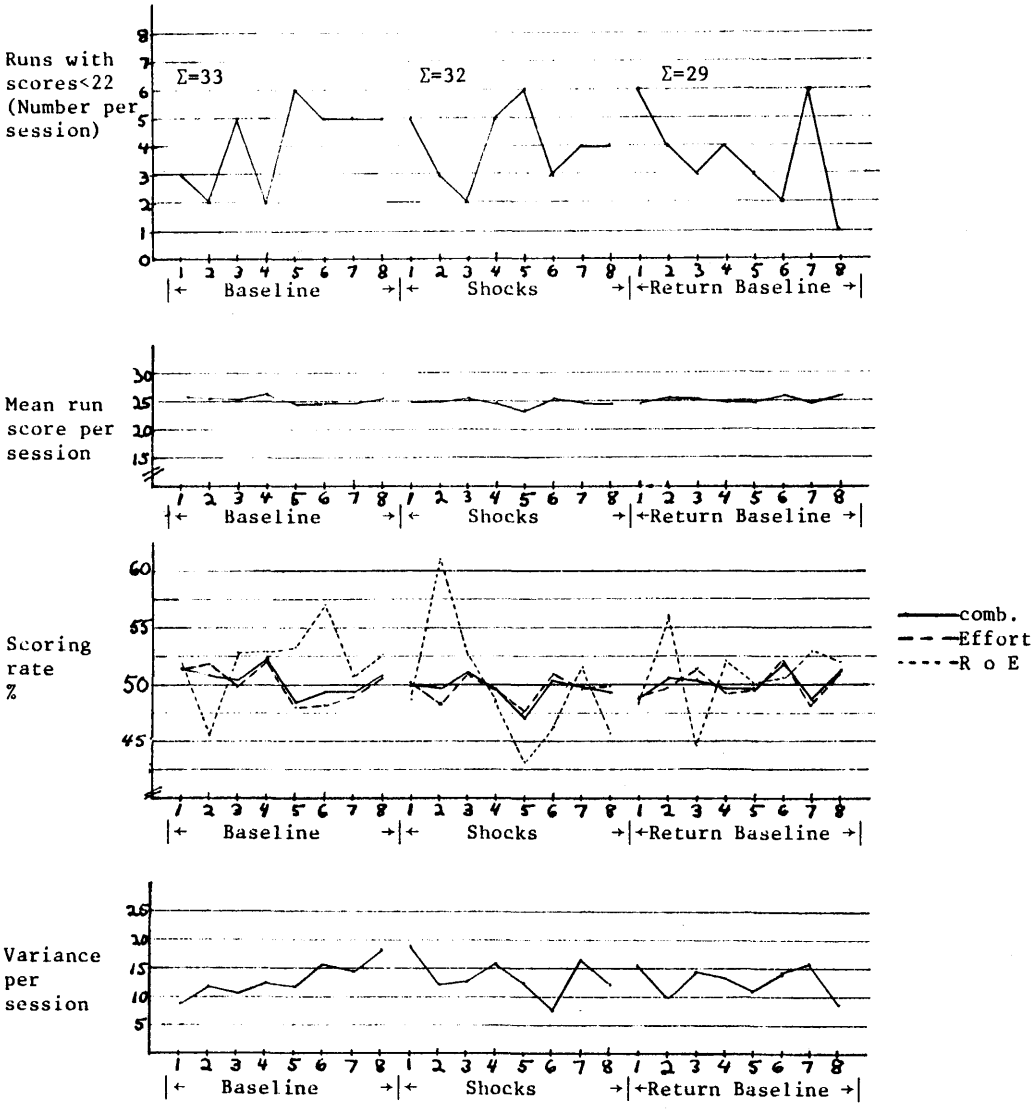


Fig. 8 - Confirmation results for subject M.J.

scores, failed to alter scoring patterns in four representative subjects in a PK task. The two interpretations which may be given to this are either 1) that the shock punishment was not an adequate conditioner or 2) that PK is inherently unconditionable.

Some support for the first conclusion may be found in the fact that all subjects reported that while the shocks were decidedly unpleasant a certain amount of adaptation was inevitable and the fear of the shocks diminished with time. It is true that the same happens in normal behaviour modification but it could be argued that a still more potent conditioner is necessary for paranormal activities to be affected. Since this argument is subject to infinite regress, to put too much weight upon it may be unwise.

Elements of operant learning are present in a remarkably wide range of human behaviour, from eye-blinks to mystical experience, However, an operant schedule may be effective under just a very circumscribed set of conditions. In bio-feedback control, for instance, learning occurs only when an artificial external feedback signal is given of the behaviour in question. Furthermore, while in the sequestered environment of the laboratory miracles of physiological control can be demonstrated, many subjects are unable to retain that control (of, say, blood pressure) when subjected to the stronger competing stimuli of the outside world.

In the present experiment no direct attention was paid to the internal state of the subjects. Perhaps electric shock is positively disruptive of some state necessary to produce PK effects. For ESP many workers believe that a state of lowered cortical arousal is beneficial. If such were true for PK, then stimulating inputs would be inappropriate conditioners. However, while a few experiments do suggest that perhaps lowered critical arousal may facilitate PK it is quite clear that this is by no means a necessary condition. To cite only one example Schmidt's PK work, some of the most successful on record, has typically involved getting the subjects involved in and highly aroused by the task. If such a 'state' view is, nonetheless, adopted a more fundamental problem arises for operant conditioning. It is difficult, or perhaps impossible to find any reward or punishment, which is not stimulating by its very nature.

This leads directly to the second possible conclusion: our findings do at least raise the question of whether paranormal abilities are, in principle, conditionable. Admittedly such a view suffers from the very

same problem as trying to 'prove' the existence of ESP, i.e. the difficulty of trying to demonstrate a universal negative, but it remains the case that parapsychological research has provided no unequivocal evidence for learning or conditioning effects in psi behaviour. It cannot be ignored that the technique which failed to get results in this experiment is not just another 'good idea' but is, in fact, one of the most potent conditioners of human behaviour known and its failure in this situation may say something about the nature of psi. Can one be sure that psi can be manipulated by the familiar psychological methods? This basic assumption may need to be questioned.

Where might this lead? We shall here outline only a single possibility. A currently unfashionable view of the paranormal is that it resembles more a defect or disease than it does a sensorimotor skill. If this is so, attempts to interfere with the normal functioning of the nervous system by methods which require its integrity, may not prove very productive (cf Walter, 1965).

While no firm conclusions can come from an experiment with thoroughly negative findings this work may be of assistance in helping other investigators decide whether or not this particular area is a potentially fruitful one in which to work.

SUMMARY

In this experiment behaviour modification techniques were used in an attempt to condition PK scoring. Four subjects performed PK tests on a game-like device and during the experimental condition all runs falling below a preset criterion were punished by an electric shock. The experiment employed an ABA design common in behaviour modification research and subjects performed 8 sessions of 24 runs (50 binary trials per run) for a total of 192 runs per condition. Both a pilot and a confirmatory study were run.

No evidence of any PK conditioning was found. A number of subsidiary effects found in the exploratory analyses of the pilot study failed to be confirmed.

ACKNOWLEDGEMENTS

The authors are indebted to Dr. Ian Tierney of the Department of Psychiatry, the University of Edinburgh for his assistance in planning parts of the experiment. Special thanks are due Dr. Sybo Schouten of the Parapsychology Laboratory of the University of Utrecht for the fearless dedication to science shown by him in joining the authors as a subject in this experiment.

REFERENCES

Azrin, N.H. and Holz, W.C. 'Punishment', in W.K. Honig (ed), 'Operant behavior: Areas of research and application', New York, Appleton Century Crofts, 1966.

Beech, H.R. 'Changing man's behaviour', London, Penguin, 1969.

Broughton, R.S. 'An exploratory study on psi-based experimenter effect and subject expectancy effects', in 'Research in parapsychology 1976', Metuchen, Scarecrow Press, 1977.

Broughton, R.S. 'An experiment with the head of Jut', E.J.P., 1979, 2, 337-357.

Camstra, B. 'PK conditioning', in 'Research in parapsychology 1972', Metuchen, Scarecrow Press, 1973.

Gatlin, L.L. 'A new measure of bias in finite sequences with applications to ESP data', J. A.S.P.R., 1979, 73, 29-43.

Hilgard, E.R. and Bower, G.H. 'Theories of learning', Englewood Cliffs, N.J., Prentice Hall, 1975.

Honorton, C. 'Psi and internal attention states', in B.B. Wolman (ed), 'Handbook of Parapsychology', New York, Van Nostrand Reinhold, 1977.

Kanfer, F.H. and Phillips, J.S. 'Learning foundations of behavior therapy', New York, John Wiley and Sons, 1970.

Kennedy, J.E. and Taddonio, J.L. 'Experimenter effects in parapsychological research', J.o.P., 1976, 40, 1-33.

Millar, B. 'The observational theories: A primer', E.J.P., 1978, 2, 304-332.

Millar, B. 'The distribution of psi', E.J.P., 1979, 3, 78-110.

O'Brien, J.T. 'An examination of the checker effect', in 'Research in parapsychology 1978', Metuchen, Scarecrow Press, 1979.

Paul, G.L. 'Behavior modification research: Design and tactics', in C. Franks (ed), 'Behavior Therapy: Appraisal and status', New York, McGraw Hill, 1969.

Pratt, J.G. 'Prologue to a debate: Some assumptions relevant to research in parapsychology', J. A.S.P.R., 1978, 72, 127-139.

Sargent, C.L. 'Repeatable significance and the significance of repeatability', Paper read to the Third Annual S.P.R. Conference, 2-4 April 1979, Edinburgh, 1979. (see also this issue)

Solomon, R.L. 'Punishment', Amer. Psychologist, 1964, 19, 239-253.

Stanford, R.G. 'Experimental Psychokinesis: A review from diverse perspectives', In B.B. Wolman (ed) 'Handbook of Parapsychology', New York, Van Nostrand Reinhold, 1977. (a)

Stanford, R.G. 'The application of learning theory to ESP performance: a review of Dr. C.T.Tart's monograph', J. A.S.P.R., 1977, 71, 55-80. (b)

Tart, C.T. 'Card guessing tests: Learning paradigm or extinction paradigm?', J. A.S.P.R., 1966, 60, 46-55.

Tart, C.T. 'The application of learning theory to ESP performance', New York, Parapsychology Foundation, 1975.

Tart, C.T., Palmer, J. and Redington, D.J. 'Effects of immediate feedback on ESP performance: a second study', J. A.S.P.R., 1979, 73, 151-165.

Thouless, R.H. 'Some experiments on PK effects in coin spinning', Proc. S.P.R., 1945, 47, 277-280.

Walter, W.G. 'The living brain', Penguin Books, 1965, 169-170.

Wheeler, J.A. 'Not consciousness but the distinction between the probe and the probed as central to the elemental quantum act of observation', Paper read to Annual convention of the AAAS, January, 1979.

White, R.A. 'The influence of persons other than the experimenter on the subject's scores in psi experiments', J. A.S.P.R., 1976, 70, 133-166. (a)

White, R.A. 'The limits of experimenter influence on psi test results: Can any be set?', J. A.S.P.R., 1976, 70, 333-370. (b)

THE POTENTIAL OBSERVER EFFECT OR THE MYSTERY OF IRREPRODUCEABILITY

Dick J. Bierman and Joop M. Houtkooper

The apparent violation by some of the psi phenomena of the rules governing the ordinary development of time (where causes precede their effects) seems to fit nicely into the so called space-time independency of the observational theories (Millar, 1978; Schmidt, 1978; Walker, 1978). A closer examination of this property in those theories reveals some problems which are so serious that they undermine the credibility of the observational theories to a significant degree.

Space-time independency is generally operationalized as follows: The result of a psi experiment is independent from the temporal and spatial distance between the psi source(s) and the psi event. In the observational theories the psi source is the observer of the experimental results. In a truly space-time independent model one does not discriminate between first and subsequent observers. All these observers contribute an amount corresponding to their psi strength. It has been shown (Houtkooper, 1977; Bierman, 1977) that the implication of this space-time independent model (under certain mathematical assumptions with regard to the additivity of psi sources) is an increasing divergence of the experimental results with an increasing number of observers. An infinite number of observers would result in

---

Note: This paper fulfils the publication policy of this journal.

experiments which show either pure hitting or pure missing.

This problem is called the 'divergence problem' and can be stated in a somewhat different way: Any effect found in psi experiments might be explained by assumptions about some future observers (e.g. some super psychics). In this form it illustrates clearly the fact that space-time independency is a feature which undermines the theory. A theory with such an explanatory power is just too good. It not only explains irreproduceability but it tells us that it is impossible ever to obtain reproduceable results because the experimenter has no control over the possible future psi sources.

The problems of space-time independence are not insoluble in principle. This has been shown by the development of a hierarchical model including the absorbance of psi (Houtkooper et al, 1980). In this curtailed model space-time independence is maintained for the length of time intervals but not for the time order in which subsequent observations take place. This order is used to derive levels of observation, each level being accompanied by a partial 'absorption' of the psi coming from the higher levels. Although this may well be a tenable position we must mention that time-order of observers is not generally preserved under relativistic transformation. Thus time-order is dependent upon the standpoint of the observer and therefore contains some arbitrariness. In the following we try therefore to preserve time-independence in the original uncurtailed way.

The solution of the forementioned problems then is to assume another mathematical formalism to describe the interaction of multiple psi sources. Instead of addition, Miller and Hartwell (1978) proposed an averaging rule. Indeed the divergence problem is solved as well as the problem of possible future super psychics. These can only have a fractional and thus limited influence. However another problem arises with this averaging model. If one assumes a normal distribution of psi sources among the population it follows directly that for a large number of observers it is impossible to get any psi effect. Although this might be more in accordance with reality than the 100% scoring which is predicted under the original assumption neither is it very satisfactory.

To allow for occasional psi effects a new interpretation is given by Millar and Hartwell to the general idea of the observational theories. Not all the observers, present and future, are contributing with the

same weight. On the contrary: an observer contributes with a weight factor according to the probability that he was to become an actual observer. So the potentiality of being an observer plays a role. An illiterate chinese has virtually no chance ever to observe the data. Therefore he doesn't play a role; even when it actually occurs that he observes those results. On the other hand, a careful reader of the parapsychological journals does contribute even when he actually misses that special article with those figures.

The best way to illustrate the potential observer model and its ambiguities is to consider the generally observed phenomenon that the filming of poltergeist events tends to block the occurrence of these phenomena (Mathews and Solfvin, 1976). This occurs also in cases where the introduction of new people on itself did not result in a decrease of the incidence of events. The potential observer model explains this effect by assuming that the act of filming drastically increases the number of potential observers whereas the number of actual observers of this uninteresting film may be close to zero.

A closer examination of this example brings to light two possible criticisms. In the first place the potentialities are obviously developing in time. The potentiality of a large number of observers for that film drops after it turns out that the moment the camera was turned on the phenomena ceased. So the question is when do we determine these potentialities. The second criticism is that the actual potentialities can in practice only be estimated within a certain framework of imperfect knowledge about the state of the universe. This imperfection may set the application of potentiality in practical cases on an uncertain footing. The solution of this problem is on the one hand that as long as practical, imperfect estimates produce reasonably good predictions, they should be used. On the other hand it is easily possible to construct an experimental design in which part of the potentialities are well defined. With regard to the moment of determination of the values of these potentialities we assume that they are determined on the moment of first observation. (That is in our experiment: the observation by the subject.) The effect of potentiality can thus be tested as follows:

In the present experiment we have manipulated the number of potential observers of the individual random events from 1 to about ten. We expect from this manipulation a differential effect.

Furthermore we have lowered the potentiality of the results to be

observed by the experimenters and readers of this paper by randomly destroying data from 50% in the pilot to 75% in the confirmatory study. Through this procedure we hoped to screen off the possible differential effect brought about by psi influences of the experimenters and the readers of this paper (Read on, dear subjects).

THE PILOT STUDY

Experimental set up and procedure

The 10 subjects volunteered to participate. Each subject did 20 runs of PK on a binary random number generator (RNG). A run consisted of 512 trials at a rate of approximately 25 trials per second. During the run the performance of the RNG was displayed graphically in real time. At the end of the run the subjects got a single line message on the display concerning the total run result. For runs with a below chance score this feedback was always the same and not discouraging. However for above chance runs different texts were available according to the scoring level. This feedback was in general somewhat over optimistic.

Auditory feedback was given as a buzz sounding at the occurrence of a series of 6 or more subsequent hits (called bursts). Each session started with a 10 minute instruction by one of the two experimenters who were both present in the instructionroom. Then a testrun was done with a random target direction, after which the target direction for the first 10 experimental runs was chosen by the subject. The subject was left alone during these runs. Between run 10 and 11 there was a break of about 10 minutes. The runs 11 - 20 had their target direction reversed from the first 10 runs. All the data were stored on magnetic disk.

The manipulation of number of potential observers. COPY versus NONCOPY condition

It was explained to the subject that after each run he could rest for a moment while the terminal produced a hard copy of the final graphic representation of the RNG performance during the run. This copying produced some noise. After the testrun the hard copy produced

was taken out of the terminal and used to explain the picture to the subject.

However the subjects were cheated to some degree with this story. Although after each run the terminal seemed to produce a hard copy, it only produced a real hard copy in runs 2,3,6,7,10,11, etc.. (ABBA scheme) After the experiment was over the subject received these 10 hard copies with the request to show them to friends and family. Thus the number of potential observers for these copied runs were larger than the number of potential observers of the non copied runs. These latter were only to be observed by the subject.

Further treatment of the data

The destruction

It was planned that after the experimental runs but before the start of the analyses 50% of the data would be destroyed. This reduces the probability for each data item, which was observed by the subject, to be observed by a future observer (including the experimenter-analyzer) to 50%. Thus reducing the 'influence' of these future observers following the POE model. The procedure used for this destruction was RNG based randomly per subject and condition. Thus after the procedure we, as analyzers, were left with the data of 5 runs in the COPY and 5 runs in the NONCOPY condition per subject.

The analyzers

In a previous experiment one of the authors had found a dramatic effect of the analyzer on the results (Weiner and Bierman, 1978). Therefore it was decided to divide the data of the 10 subjects randomly over the two experimenters. Each of the experimenters got 5 subjects to analyze. It was felt that if the predicted preferential effect should be present in both data sets this would strongly support the reality of this effect. On the other hand if differences between the two data sets should occur this might be interpreted as disadvantageous for the proposed potential observer model. The potentiality for both data sets to be analyzed by each of the 2

analyzers being equally decreased by the destruction procedure.

Hypotheses

a) The main hypothesis was a differential effect of the DIRECT PK scores between the COPY and NONCOPY condition. (DIRECT(ed) PK occurs when the random process is consistently 'influenced' in the target direction (hitting) or in the non-target direction (missing).) With regard to the direction of the predicted difference it was felt that the averaging rule would imply that the scores in the COPY condition should be closer to chance than in the NONCOPY condition due to the larger number of potential observers in the COPY condition.

b) Secondary hypotheses were on differential effects of the RANDOM PK scores between the conditions. (RANDOM PK occurs when the 'influence' on the random process is fluctuating between missing and hitting.) These hypotheses were to be tested with, as a unit of observation, firstly the run and secondly 32-trial segments.

c) Exploratory analyses were planned on the burstscores.

RESULTS

All the results are presented in two parts. In the first part the effect of the COPY condition is analyzed. That is the effect of the number of potential observers. This analysis is done by a one sample t-test on the difference score (DIFF) between the NONCOPY and COPY condition. (So DIFF=NONCOPY-COPY). Furthermore a one sample t-test is done on the total score (SUM). (SUM=NONCOPY+COPY).

In the second part the total data set is split into two parts, each containing the data of 5 subjects. The data of subjects 1 - 5 have been analyzed by JMH while the data of subjects 6 - 10 have been analyzed by DJB. A possible analyzer effect is tested for by a two sample t-test using the data analyzed by JMH as the first and the data analyzed by DJB as the second sample. (Positive td-values mean a larger mean-value in the first than in the second sample)

Direct PK scores

The deviation of the number of random events from chance are expressed in z-scores. These scores are given per subject in table 1. Remember that the COPY and NONCOPY data are the basic data and that the data of difference and sum scores are derived from them.

TABLE 1
Normalized DIRECT PK scores for the two conditions and normalized differential and sum scores

Subjects	COPY	NONCOPY	DIFF	SUM
1 ARJ	-.119	-.356	-.168	-.335
2 WOU	.791	-.474	-.894	.224
3 WER	.909	1.225	.223	1.509
4 MARL	-1.107	.198	.923	-.643
5 MARR	-.672	-2.055	-.978	-1.929
6 JEA	.830	-1.463	-1.621	-.447
7 BOG	1.265	-.356	-1.146	.643
8 GOV	1.423	.514	-.643	1.370
9 JOS	-.435	-.356	.056	-.559
10 MAR	-.119	-.079	.028	.388
All subjects	.875	-1.013	-1.333	-.096

Condition effect in DIRECT PK scores

A one sample t-test yields the following results:
DIFF: t=-1.745, df=9, p=.096 two-tailed.
SUM: t=.095, df=9, n.s.
There is a marginal differential effect opposite to the direction predicted. The reason for a marginal significant t-value while the

corresponding z-score is far from significant is the reduced experimental variance in the normalized DIFF-scores (sd=.765).

Analyzer effect in DIRECT PK scores

Two sample t-tests using the results of subjects 1 - 5 (analyzer JMH) as the first and the results of the subjects 6 - 10 (analyzer DJB) as the second sample yield the following results:
DIFF: td=1.006, df=8, n.s.
SUM: td=-.772, df=8, n.s.
Although the DIFF-data of analyzer DJB are marginally significant in themselves (t=-2.03, df=4) there is no significant difference between the data of the two analyzers.

RANDOM PK over runs

The total scores which are given in table 1 are the sum of the 5 runscores in each condition. The variance in the distribution of these runscores is expressed as a chi-square statistic. If some of the runs show psi missing and others show hitting the DIRECT PK score in table 1 might very well be at chance. However the chi-square statistic in this case would be significantly increased.

Condition effect in RANDOM PK over runs

A one sample t-test can not be applied directly on the data from the DIFF and SUM columns because the distribution of chi-square scores is far from normal for 5 degrees of freedom. Therefore the data of the COPY and NONCOPY condition were first transformed to z-scores according to Wilson and Hilferty. The resulting t-tests give the following results:
DIFF: t=-.598, df=9, n.s.
SUM: t=-.800, df=9, n.s.

TABLE 2
RANDOM PK over runs

Subjects	COPY	NONCOPY	DIFF	SUM
1 ARJ	2.867	3.930	-.309	-.528
2 WOU	1.625	2.359	-.305	-1.493
3 WER	4.430	11.539	- 1.211	1.244
4 MARL	10.250	6.273	.643	1.468
5 MARR	2.508	7.422	-1.161	.081
6 JEA	1.164	3.258	-.844	-1.439
7 BOG	3.703	0.336	1.671	-2.012
8 GOV	6.031	4.086	-.435	.296
9 JOS	.805	3.461	1.142	-1.618
10 MAR	7.273	5.156	.412	.779

Analyzer effect in RANDOM PK scores

Again the data from subjects 1 - 5 were compared with those from subjects 6 - 10 using a two-sample t-test. This yielded the following results:
DIFF: td=.943, df=8, n.s.
SUM: td=1.214, df=8, n.s.

RANDOM PK over segments

RANDOM PK can also occur as alternating hitting and missing within a run. Then the runscore might be at chance while extra chance variance occurs. To test for this phenomenon one has to divide the run into segments. The problem is that about any choice about the length of the segment is rather arbitrary. We decided to divide the run in 16 segments of 32 trials. The raw results of the segment variance are

given in table 3.

TABLE 3
RANDOM PK over segments

Subjects	COPY	NONCOPY	DIFF	SUM
1 ARJ	84.4	73.6	.607	-.052
2 WOU	75.8	82.5	-.380	-.029
3 WER	85.4	93.4	-.416	1.078
4 MARL	84.3	81.9	.129	.412
5 MARR	84.4	113.0	-1.419	1.974
6 JEA	77.6	77.4	.014	-.208
7 BOG	75.8	73.1	.154	-.563
8 GOV	95.5	78.1	.922	.785
9 JOS	82.1	60.8	1.286	-.976
10 MAR	93.4	75.2	.980	.513

Note: Chi-squares with df=80; DIFF and SUM converted to z-scores

Condition effect in RANDOM PK over segments

Although the chi-square values are about normally distributed we applied the same transformation to obtain z-scores as was done for the RANDOM PK scores over the runs. The one sample t-test on the DIFF and the SUM column yielded the following results:
DIFF: t=-.736, df=9, n.s.
SUM: t=1.087, df=9, n.s.

Analyzer effect in RANDOM PK over segments

A two sample t-test taking the data of subjects 1 - 5 as the first and the data of subjects 6 - 10 as the second sample yielded the following results:

DIFF: td=2.307, df=8, p=.050 two-tailed
SUM: td=1.519, df=8, n.s.

RANDOM PK within a run for other segment

The analysis to check for within run variance effects which was initially done for a segment length of 32 trials was repeated for segment lengths of 2,4,8,16,64,128 and 256 trials (see table 4). For the variance in the SUM scores there is a pretty consistent analyzer difference, which is not so remarkable as the figures are not quite independent. However for the variance in the DIFF scores there is a strange exception if the run is divided into segments of 128 and to a lesser degree of 256 trials. There the analyzer difference is reversed. Note that the results of this quite exploratory analysis are given in terms of z-scores and that the z-score for the analyzer difference in the variance of 32 trial segment DIFF-scores is far below significance level, while in the t-test for that case a significant effect was found. Again this can only be explained by a reduced variance in the scores. (Hence a reduced variance in the variance scores.)

Burst scores

For each run the number of events of 6 or more subsequent misses was subtracted from the number of events of 6 or more subsequent hits. This yielded a burst score. The COPY and NONCOPY burst scores were combined into table 5.

Condition effect in burst scores

Single sample t-tests yielded the following results:
DIFF: t=.187, df=9, n.s.
SUM: t=-.552, df=9, n.s.

TABLE 4
Within run RANDOM PK effects for different segment lengths and analyzer differences (Zd)

Segment length	DIFF			SUM		
	Ss 6-10	Ss 1-5	Zd	Ss 6-10	Ss 1-5	Zd
2	-30	84	.502	12806	12944	.608
4	-145	42	1.475	6449	6420	-.181
8	-56.5	17	.648	3193	3286	.821
16	-61.7	44	1.310	1585	1695	1.368
32	-59.6	30.8	1.587	789	859	1.217
64	-31.8	15.1	1.186	376	430	1.351
128	13.0	-14.6	-.986	175	216	1.463
256	13.6	3.8	-.602	77	119	2.104
512 (is run)	-2.7	9.8	.903	35.3	53	1.355

Analyzer effect in burst scores

Two sample t-tests between the data of the first 5 and the second 5 subjects yielded:
DIFF: td=-1.677, df=8, p=.132 two-tailed
SUM: td=.047, df=8, n.s.

Exploratory analysis burst scores

None of the specific burst lengths differed significantly from chance. An analysis of the burst score distributions as they occurred in the various parts of the runs did not yield any consistent result whatsoever.

Explorations: The first segment analysis

TABLE 5
Burst scores

Subjects	COPY	NONCOPY	DIFF	SUM
1 ARJ	-3	-6	3	-9
2 WOU	1	3	-2	4
3 WER	-1	10	-11	9
4 MARL	-7	-1	-6	-8
5 MARR	-2	0	-2	-2
6 JEA	8	-3	11	5
7 BOG	-1	-2	1	-3
8 GOV	1	3	-2	4
9 JOS	-4	1	-5	-3
10 MAR	5	-13	18	-8
Totals	-3	-8	5	-11

The analyses for RANDOM PK for different segment lengths yielded as a spin off distinct results for the scoring in the very first segment of the run. To our surprise it turned out that the data of the subjects 1 - 5 and those of subjects 6 - 10 were significantly different in this first segment.(For the first 32 trials of the runs the difference in DIRECT PK scoring between the two groups has a Zd=2.44. The scoring rate in the first segments of the data of subjects 1 - 5 (analyzer JMH) was 48.81% while the scoring rate in the corresponding data of subjects 6 - 10 (analyzer DJB) was 54.31%. A very post hoc and more detailed analyses showed this analyzer effect to be most pronounced in the first 11 trials (Zd=3,196)).

CONCLUSIONS

The marginally significant differential effect between the NONCOPY

and COPY condition was considered to be encouraging irrespective of the fact that the direction was opposed to the one predicted. It was decided to change our main hypothesis in accordance with this finding.

As the effect was mainly present in the data of subjects 6 - 10, data which were analyzed by DJB, it was felt necessary to raise the destruction rate in the confirmatory experiment in order to screen off analyzer effects more effectively. This was stressed the more because in the other analyses of these pilot data indications were found for analyzer effects. For instance in the burst DIFF-scores and very much so in the RANDOM PK DIFF-scores. The auditory feedback of events were 6 or more subsequent hits are produced did not yield any remarkable results with regard to the number of these events. The first segment analyzer effect which was discovered accidentaly is very post hoc in the first place and rather difficult to interpret in the second place. The only speculation which might be put forward is that the subjects' psi was working in the first few trials in a form that was 'structured' by the analyzers psi.

To conclude our evaluation of the pilot, an overall estimate is given with regard to the support for the Potential Observer Effect. In all our analyses here there are two effects of prime interest, firstly the condition effect hypothesized, being in accordance with the POE model: secondly at variance with this model was the analyzer effect. To combine different aspects, we took the sum of the squared t-values for the condition effect (DIFF) and for the analyzer effect (also for DIFF). This results in S=3.98 for the condition were 4 is expected in contrast with the S=10.04 for the analyzer effect. The constructed measure looks like a chi-square but the constituent t's are slightly interdependent. It can only be used to get an idea about the arguments pro and con the existence of POE, with the negative view as yet prevailing.

THE CONFIRMATORY EXPERIMENT

Set up and procedural changes

Two major changes were made with respect to the set up and procedure of the pilot study. As the analyzer effects in the pilot tended to be stronger than the 'potential observer' condition effect it was felt

necessary to raise the data destruction rate from 50% to 75%. According to the POE model this would result in a reduced influence of future observers including the analyzers. To maintain the total amount of final data (after destruction) the number of runs was increased from 20 to 24 and the number of subjects from 10 to 20. In fact this results in a slightly increased data set. (In total 60 runs in each condition versus 50 in the pilot study).

The second change was a change in procedure. The experimenters who had jointly handled the subjects in the pilot decided to work alone in the confirmatory study. This change was purely based upon the fact that the handling of twice the number of subjects would be too time consuming. So each of the experimenters handled 10 subjects in the confirmatory study. It was decided beforehand which experimenter was to work on which day.

Hypotheses in the confirmatory study

The hypotheses are deduced from the marginally significant findings in the pilot.

a) The main hypothesis was again a differential effect in the DIRECT PK score between COPY and NONCOPY condition. The scoring rate in the COPY condition was expected to be higher than the scoring rate in the NONCOPY condition.

b) An analyzer effect is predicted with regard to RANDOM PK over segments, in the difference between NONCOPY and COPY conditions.

c) An analyzer effect is predicted with regard to the PK BURST scores, also in the difference between NONCOPY and COPY condition.

These latter effects are hypothesized although the set up was changed in order to eliminate the future observer 'influence'.

Exploratory analyses

All the analyses which were done in the pilot and which did not yield hypotheses in this confirmatory study are repeated on an

exploratory basis. The idea behind this approach is that we actually did change the set up of the pilot study to suppress analyzer effects and to increase condition effects. The goal of these exploratory analyses is to see whether the expected effect of an enhanced destruction rate is indeed present.

The significant first segment PK analyzer effect in the pilot is considered to be too post hoc to base a hypothesis on.

The change in the procedure with respect to subject handling opens a nice possibility to test for experimenter/handling in contrast to experimenter/psi effects. This analysis is exploratory and we do not expect to find an effect because in our opinion most of the experimenter effects are of the experimenter/psi type.

RESULTS

Like in the pilot study all the results will be presented in such a way that a direct comparison can be made between the effects caused by the COPY condition and the analyzer effects. In this confirmatory part the results of a third effect are given simultaneously, i.e. the effect of the experimenter/handling. The analyzers performed the appropriate t-tests on the randomly divided data sets as usual. This resulted in a confusingly large number of t-values. Therefore it was decided to give the results in terms of an analysis of variance. The analyzer and the experimenter/handling are the two level grouping factors and the COPY and NONCOPY condition scores are the dependent (repeated measure) variables.

DIRECT PK

The DIRECT PK scores are given in table 6 as the total deviation from MCE for the 3 runs per condition and per subject.

Note that the main effects given here correspond to the effects in the SUM score which was used in the pilot study while the condition effects given in this table corresponds to the effects in the DIFF score.

TABLE 6
DIRECT PK scores

Analyzer DJB Subject	COPY	NONCOPY	EXP	Analyzer JMH Subject	COPY	NONCOPY	EXP
ANN	0	13	JMH	LAU	-11	13	JMH
PIM	-12	7	JMH	GER	26	5	JMH
OPH	-16	-21	JMH	VAN	24	-5	JMH
JOH	-26	-6	JMH	MAR	-12	0	JMH
FRA	-11	-10	JMH	PETv	12	-44	JMH
KIK	7	-1	DJB	WIL	-9	-38	DJB
HANs	27	1	DJB	LIL	20	-4	DJB
PETs	-15	6	DJB	JOS	25	28	DJB
HANi	13	-10	DJB	PSU	8	12	DJB
JAA	14	38	DJB	DOR	33	2	DJB

Analysis of variance

MAIN effects Source	F	p	CONDITION effects Source	F	p
mean	.451	.511	COP	1.213	.287
An	.496	.491	COP.An	3.296	.088
Ex	3.496	.080	COP.Ex	.422	.516
An.Ex	.567	.463	COP.An.Ex	.276	.606

Hypothesis 'a' on the condition effect in the DIRECT PK data cannot be confirmed as the COP effect is too small with p=.287. The F-test entails a two-tailed test. However, applying a one sample t-test on the DIFF scores as we did in the pilot study gives t=-1.073, df=19. This difference is in the same direction as in the pilot and quite interesting is the contribution of the data of analyzer JMH to this result (t=-1.901, df=9, p=.073).

It should be remarked that the marginal Ex and the An effect are caused by the significant contributions from the COPY scores. (An analysis of variance on these scores yields an experimenter effect with p=.044 and an analyzer effect with p=.065)

RANDOM PK over runs

Table 7 gives the RANDOM PK scores over runs for both conditions. These are chi-squares for 3 degrees of freedom. After transformation to z-scores according to Wilson and Hilferty these RANDOM PK scores are fed into the analysis of variance also presented in this table.

Hypothesis 'b' on the analyzer effect in the COPY-NONCOPY differences between RANDOM PK over run scores is represented by the COP.An interaction in table 7. The non significant p-value of .542 is fairly middle of the road.

RANDOM PK over segments

As in the pilot study the runs were divided into 16 segments of 32 trials to check for RANDOM PK within the runs. A chi-square is calculated per run (df=16) and these scores are summed over the three runs for each subject in each condition. This yields the within run RANDOM PK scores which are given in table 8 as chi-squares with df=48. The analysis of variance applied to the transformed scores is also presented in this table.

The significant COP.Ex interaction is mainly caused by the below chance RANDOM PK scores of the An(JMH).Ex(DJB) group of subjects in the COPY condition (191.1 total, with z=-2.373; p<.02 two-tailed).

TABLE 7
RANDOM PK scores over runs

Analyzer DJB				Analyzer JMH			
Subject	COPY	NONCOPY	Exp	Subject	COPY	NONCOPY	Exp
ANN	10.516	6.289	JMH	LAU	1.211	0.445	JMH
PIM	0.984	1.023	JMH	GER	3.781	8.856	JMH
OPH	1.172	2.133	JMH	VAN	3.578	1.102	JMH
JOH	4.531	0.703	JMH	MAR	1.047	1.703	JMH
FRA	2.820	2.703	JMH	PETv	4.188	5.625	JMH
KIK	1.164	2.258	DJB	WIL	0.227	4.156	DJB
HANs	3.977	5.570	DJB	LIL	3.141	0.063	DJB
PETs	0.883	0.766	DJB	JOS	3.914	2.125	DJB
HANi	0.789	1.391	DJB	PSU	3.922	0.703	DJB
JAA	1.703	4.156	DJB	DOR	2.898	0.953	DJB

Analysis of variance

Main effects			Condition effects		
Source	F	p	Source	F	p
mean	.389	.542	COP	.358	.558
An	.063	.804	COP.An	.508	.486
Ex	1.133	.303	COP.Ex	.025	.875
An.Ex	.031	.863	COP.An.Ex	2.484	.135

TABLE 8
RANDOM PK over segments

Analyzer DJB Subject	COPY	NONCOPY	Exp	Analyzer JMH Subject	COPY	NONCOPY	Exp
ANN	60.00	57.63	JMH	LAU	48.88	40.13	JMH
PIM	41.75	38.13	JMH	GER	46.[illegible]	39.63	JMH
OPH	48.25	52.63	JMH	VAN	47.25	43.38	JMH
JOH	53.00	45.50	JMH	MAR	42.75	45.25	JMH
FRA	52.63	37.75	JMH	PETv	43.50	62.25	JMH
KIK	40.63	43.13	DJB	WIL	33.63	45.25	DJB
HANs	46.13	53.63	DJB	LIL	28.00	60.75	DJB
PETs	57.88	73.00	DJB	JOS	47.12	49.00	DJB
HANi	45.63	35.50	DJB	PSU	40.00	52.75	DJB
JAA	35.50	49.75	DJB	DOR	42.38	41.50	DJB

Analysis of variance

Main effects Source	F	p	Condition effects Source	F	p
mean	.654	.431	COP	1.702	.210
An	1.326	.266	COP.An	1.533	.233
Ex	.333	.572	COP.Ex	5.360	.034
An.Ex	.054	.820	COP.An.Ex	.053	.820

Burst scores

A burst is a series of 6 or more events in the same direction. In case of such a burst in the target direction an extra audio feedback was given, as in the pilot study. The number of bursts in the target direction minus the number of bursts in the nontarget direction is given in table 9, together with the results of the analysis of variance.

Hypothesis 'c' cannot be confirmed as the COP.An interaction is too small, with p=.277. Moreover the small effect present has the opposite direction compared with the effect in the pilot. An interesting occurrence is the unpredicted and unwanted Ex effect with its main contribution from experimenter DJB, who elicited 62 more target bursts than non target ones from his subjects.

First segment effects

In the pilot study a post hoc effect was found with respect to the DIRECT PK scores in the very first beginning of the runs. In table 10 these scores are given for the confirmation study. Each score is the sumscore of the first segment of three runs. The analysis of variance on these data is also presented in this table.

The significant analyzer effect in these first segments of the run data, An(DJB) obtaining hitting with a rate of 51.72% and An(JMH) obtaining missing with a rate of 47.92% replicates the rather curious finding in the pilot study. Post hoc we calculated two sample t-tests for this analyzer effect and obtained for the pilot td=3.163, df=8 and in the confirmation td=2.759, df=18. The p-values are respectively .0133 and .0129 two-tailed.

DISCUSSION

After seeing our best choice of three hypotheses unconfirmed, while the non hypothesized first segment analyzer effect replicated like a planetary orbit, we feel a bit disappointed. As far as the true purpose of this research is concerned, the probing of the main ideas

TABLE 9
Burst scores

Analyzer DJB Subject	COPY	NONCOPY	Exp	Analyzer JMH Subject	COPY	NONCOPY	Exp
ANN	-1	0	JMH	LAU	-5	2	JMH
PIM	-4	3	JMH	GER	10	0	JMH
OPH	-8	-3	JMH	VAN	5	-5	JMH
JOH	-11	-1	JMH	MAR	-5	1	JMH
FRA	0	-6	JMH	PET	7	-8	JMH
KIK	3	-3	DJB	WIL	-7	-8	DJB
HANs	9	1	DJB	LIL	9	5	DJB
PET	1	5	DJB	JOS	12	0	DJB
HAN	5	-1	DJB	PSU	0	10	DJB
JAA	3	13	DJB	DOR	3	2	DJB

Analysis of variance

Main effects Source	F	p	Condition effects Source	F	p
mean	.927	.350	COP	.272	.609
An	.451	.512	COP.An	1.268	.277
Ex	7.051	.017	COP.Ex	.061	.808
An.Ex	1.574	.228	COP.An.Ex	1.033	.325

TABLE 10
First segment DIRECT PK scores

Analyzer DJB Subject	COPY	NONCOPY	Exp	Analyzer JMH Subject	COPY	NONCOPY	Exp
ANN	-4	-3	JMH	LAU	-2	0	JMH
PIM	1	0	JMH	GER	4	-3	JMH
OPH	-2	6	JMH	VAN	-4	-6	JMH
JOH	8	10	JMH	MAR	-5	7	JMH
FRA	13	-5	JMH	PET	4	-10	JMH
KIK	6	3	DJB	WIL	-3	-3	DJB
HANs	8	-4	DJB	LIL	2	-4	DJB
PET	6	-7	DJB	JOS	-5	-1	DJB
HAN	2	-3	DJB	PSU	3	-3	DJB
JAA	-1	-1	DJB	DOR	-7	-4	DJB

Analysis of variance

Main effects Source	F	p	Condition effects Source	F	p
mean	.065	.801	COP	2.461	.136
An	7.134	.017	COP.An	.593	.452
Ex	.837	.374	COP.Ex	.359	.558
An.Ex	.033	.857	COP.An.Ex	.684	.420

in the POE model, the condition effect at first occurred at marginal significance but opposite to the expected direction. In the confirmation the direction persisted, but at a level far below significance. Some encouragement can be obtained from looking at the results per analyzer, but no consistent picture develops.

The analyzer effects which were unwanted by the analyzers entered the picture in the pilot in the analyses of RANDOM PK effects. In the confirmation these effects disappeared from the RANDOM PK scores and returned marginally in the DIRECT PK scores. Exactly where they were not explicitly predicted. This is topped up by the strong analyzer effect in the first segments of the runs. Here no hypothesis was formulated because we thought that this effect in the pilot was too post hoc. There seems to be a kind of systematic avoidance of confirmation.

Our rough measure of overall comparison between condition effects and analyzer effects, in this case operationalized as the sum of F-scores, results in 6.006 for the overall condition effect and 7.198 for the overall analyzer effect. Both values appear to be slightly above chance (about 5) but they hardly favor one effect over the other. Optimistically one might say that the apparent relative decrease of analyzer effects is caused by the increased destruction rate which directly favors the POE model. There are several indications for experimenter effects which might be explained by the different approach to the subjects of the two experimenters. But what to say about the significant Experimenter Condition interaction?

A recent experiment in which destruction rate within a run was manipulated yielded results in which analyzer effects were interacting with the destruction variable. However the form of the interaction was not in line with the POE model (Bierman and Weiner, 1980). Analyzer effects which occur in randomly split data are in conflict with the POE model anyway because the potentiality of each data item to be observed by each of the two analyzers is 50% independent from the actual observation.

Therefore we tend to reject the POE model although we feel that the concept of potentiality might still be a useful one, especially to describe systems before the first observation. However after the first observation conbination of psi might be governed by temporal order of observation (Houtkooper et al, 1980; Bierman, 1980).

A final remark should be made with regard to the general procedure in parapsychological experiments. In this study we followed the Pilot-Confirmation procedure. This procedure can be seen as a tool to get a measure for internal validity. If predictions based upon the pilot data are confirmed than the researcher can be more confident that his significant finding is not significant by chance.

In the course of research history a strange custom developed never to predict a scoring direction. If for instance boys scored better than girls in the pilot study the hypothesis for the confirmation was simply a significant sex effect. And indeed often directions did reverse from pilot to confirmation.

The analysis procedure in which the data are randomly divided into two parts which are compared can also be seen as a method to measure internal consistency. Like in the pilot-confirmation approach a correspondence between the two data sets will strengthen the conclusions. A significant difference between the two sets however will not be interpreted in the way it is in the pilot-confirmation paradigm. If one analyzer gets significantly better results from boys than from girls while the other analyzer finds the reverse, the conclusion will be that this difference is more related to the analyzers than to the sex variable. However people using the pilot-confirmation paradigm often reason that the experimental situation is able to elicit psi in a different way for boys and girls but the direction is not controlled.

It will not surprise the reader that we feel that such reasoning could be very misleading. We would prefer to stay on the safe side and to consider the pilot-confirmation paradigm as a special case of the analyzer split paradigm.

Therefore the use of the latter should be preferred above the use of the former although even if the data sets do show a high correspondence this could still be due to the analyzers' psi.

SUMMARY

In the present RNG based PK experiments (pilot and confirmation) 680 runs of 512 trials were done by 30 unselected subjects. Immediately following the experiments, before observation by someone else besides

the subject could take place, the data of 460 runs, randomly chosen from the total data sets were destroyed. According to a model proposed by Millar and Hartwell to remove some problems inherent to present formulations of observational theories this destruction would make the total data sets less sensitive for other psi sources than the subject.

Furthermore a predetermined number of runs was hard copied and given to the subject, thus enabling potential observation of these runs by others than the subject. A marginal effect of this COPY condition was found in the pilot. But it was not confirmed in the confirmation study. On the other hand analyzer effects were found in the pilot as well as in the confirmation. This indicates that the destruction procedure did not result in an adequate screening of the data from analyzers psi.

ACKNOWLEDGMENT

We want to thank Brian Millar for the stimulating discussions which cumulated in the present experiment.

REFERENCES

Bierman, D.J. 'Observer or experimenter effect. A fake replication', E.J.P., 1977, 2, 2, 115.

Bierman, D.J. 'A test of the potential observer theory using gifted subjects', presented at the P.A. Convention 1980.

Bierman, D.J. and Weiner, D.H. 'A preliminary study on the effect of data destruction on the influence of future observers', J.o.P., 1980, 44, in press.

Houtkooper, J.M. 'A comment on Schmidt's mathematical model', E.J.P., 1977, 2, 1, 15.

Houtkooper, J.M., Andrews, R., Ganzevles, P.C.J., and van der Sijde, P.C., 'A hierarchical model for the observational theories: A study of subject identity and feedback in repeated retroactive psychokinesis', E.J.P., 1980, 3, 3, 221-246.

Mathews, F.M. and Solfvin, G.F. 'A case of RSPK in Massachusetts', R.I.P., 1976, 219-223.

Millar, B. 'The observational theories: A primer', E.J.P., 1978, 2, 304.

Millar, B. and Hartwell, J. 'Dealing with divergence', presented at the P.A. Convention 1978.

Schmidt, H. 'A remark on the 'divergence problem'', E.J.P., 1978, 2, 2, 163.

Walker, E.H. 'The quantum mechanical theory of consciousness: A response to the critique by Chari', Indian J.P., 1, 2, 130.

Weiner, D. and Bierman, D.J. 'An observer effect in data analysis?', R.I.P., 1978, 57.

Dick J. Bierman
RIPP
Alexanderkade 1
1018 CH Amsterdam
The Netherlands

Joop M. Houtkooper
Laboratory for Psychophysiology
University of Amsterdam
le Constatijn Huygenstraat 20
Amsterdam
The Netherlands

IN DEFENSE OF INTROVERTS: A CRITICAL NOTE ON SUPPOSED PERSONALITY DIFFERENCES BETWEEN GANZFELD EXPERIMENTERS

Adrian Parker

It has long been maintained that there are clear and distinctive personality characteristics between those experimenters who obtain significant results in psi research and those who do not. However what is less recognised is that much of the evidence for this view is merely anecdotal in nature and consists of recommended ways of handling subjects based on the experimenter's own experience. The experimental evidence is either post-hoc or limited to showing that differential scoring patterns can occasionally be obtained from groups which the experimenter has handled in say an aloof, detached manner versus a warm personalised style. From the point of view of hypothesis testing it would appear that if the experimenter effect is mediated by personality the areas or traits involved are those relating to sociability and the use of social skills. (It may of course be that there is an interpersonal or interactive effect between subject and experimenter and I intend to return to this line of argument later.)

Four years ago I reported a pilot investigation (Parker, 1977) which I carried out in order to evaluate this prediction by giving a personality inventory, the 16PF, to parapsychologists who were independently classified by judges using fairly objective criteria into experimenters who consistently obtain significant results (psi-conducive) and those who usually fail to do so (psi-inhibitors). The results showed none of the predicted differences in personality factors to be statistically significant. However a study such as this has obvious flaws and weaknesses due to cross cultural problems in

standardisation. Because of the limited number of researchers who could be clearly classified by judges as psi-conducive or inhibitory, further refinements as to nationality etc. could not be made since at least at that time there were no separate standardisations of the 16PF for the various national groups. Clearly a more sophisticated study is justified before the issue is finally settled. Recently a further study has been reported by Carl Sargent (1980) using some of the original data supplied by me and supplemented by three additional 16PF profiles of British researchers he has obtained. It is with this study and the claims he makes for it that the present review is concerned.

Sargent's article begins in his customary style with a barrage of criticisms and negatively toned value judgements, which in substance amount to little more than a re-statement of the shortcomings already mentioned in my 1977 study. However to be fair he does lodge two further criticisms which are probably worth correcting before proceeding. First, in the full version presented at the 1979 P.A. Convention, Sargent claims the Form C of the 16PF to be inappropriate for use with higher educational groups such as parapsychologists. Cattell, Eber, and Tatsuoka clearly state in their 'Handbook for the 16PF' that although forms A and B are recommended for well educated groups because of the type of vocabulary used in formulating the test questions (and nothing to do with validity), ultimately the final choice of test form must be decided by time limitations (note 1). Since form C can often take 20 minutes less to complete than form A or B, it was thought that busy parapsychologists would appreciate being sent the shorter form and indeed this was considered critical in ensuring an adequate number of data returns.

Sargent's second point concerns the idea of a psi-conducive experimenter being task specific, that is one must select out experimenters for the type of research task they succeed at. Here I can follow Sargent's reasoning only so far. Clearly it is easy to believe that there is likely to be a difference between researchers who are successful with spontaneous phenomena and those who succeed with experimentation. Indeed it was an important feature of the original pilot study that only the personality profiles of experimenters who were successful or unsuccessful in psi-experimentation were used. No doubt it is also true, as Sargent says, that there are different styles of handling subjects in the test situation. A priori it does not seem implausible that this might interact with experimenter personality. However, in practice, Sargent's speculation has no factual basis. There is nothing in the

literature to suggest that researchers actually have different track records in free versus forced choice experimentation.

To turn now to Sargent's own study, by limiting the investigation to personality profiles of only Ganzfeld experimenters he claims to have confirmed two of my original hypothesis made about differences in traits and to have found two additional significant differences in comparing psi-conducive and inhibitory experimenters. Obviously, if valid this is an enormously important finding, and given the derisive nature of his criticisms, it might not be unreasonable to expect with the advantage of hindsight that his own methodology would show something of an improvement. Not so, much to my amazement in scrutinizing his raw data it appears that as well as including the flaws incumbent to my own study he goes on to add several of his own, some of a very elementary nature and sufficient to invalidate all his findings.

To begin with, Sargent adds my data to his own without any attempt at correction thereby mixing not only two forms of the 16PF (C and B) but also two separate editions (1962 and 1969). As Cattell, Eber, and Tatsuoka (1970) note several of the scales in the later edition were revised, and scores on the earlier edition are lower estimates, so it is clearly misleading to regard the two sets of data as compatible. It may also be noted that Sargent failed to make use of the now available British norms for the 16PF (Saville and Blinkhorn 1976) when scoring his additional Ganzfeld experimenters (who were all British) added to my group. This would have eliminated the problem with national differences since all the experimenters in Sargent's groups were either British or American. However it would seem from Sargent's comments that he experiences some confusion over what constitutes artifact due to nationality. In order to make a comparison between groups of psi-conducive experimenters and psi-inhibitory experimenters, it is not necessary that we have (as Sargent will) a balanced assortment of nationalities in each group. Indeed if it is shown that most psi-conducive researchers are American and most psi-inhibitory researchers are British (as my original study found) then this is an interesting although fairly well known result. However it is more important to compare each experimenter's personality with respect to the norms for his own culture (rather than merely using American norms). To choose the classic example an extraverted English personality may not be regarded in equivalent terms and evoke the same response in say the United States.

There are however more serious errors. Nowhere does Sargent state whether or not he used any objective criterion or independent method of classifying researchers as successful (conducive) or unsuccessful (inhibitory) Ganzfeld experimenters. (I used three independent assessors in the original study.) Nevertheless it may be presumed that the classification of experimenters is a fairly straightforward matter depending on their published results. However this is not the case, especially in instances where the experimenter has conducted several studies with contrasting results or where his involvement in the study is tenuous. For example Sargent quite arbitrarily classifies Rogo as a psi-inhibitory experimenter, when in fact 2 out of his 4 reported Ganzfeld studies were significant, and Stanford as a psi-conducive experimenter, when of the three Ganzfeld studies he has been involved in, two produced nonsignificant overall results and the third which produced a significant finding did not seem to involve him as an actual experimenter (note 2). Furthermore Tart is classified as a psi-inhibitory experimenter merely because his name is included along with others in the Palmer Ganzfeld study. In fact it would seem from the original report that his name was included purely out of courtesy since the research was conducted in his laboratory and that in practice he had no role as experimenter. Moreover even if one can justify the inclusion of Tart, this study actually produced overall significant psi-missing scores so it is obvious that he is either wrongly included or misclassified. (Incidentally all my judges classified Tart as 'conducive').

Equally suspect is that in order to obtain groups of psi-conducive and inhibitory experimenters, Sargent is prepared to confound studies, subjects, and experimenters with each other. Thus Harley and Sargent himself, who used each other as subject and experimenter in a joint Ganzfeld experiment (note 3), are lumped together with L. Braud, Drucker and Stanford, who have worked with volunteer subjects, in order to form a so-called psi-conducive group of experimenters. Needless to say the inclusion of Sargent's (and possibly Harley's) 16PFs as well as being somewhat anomalous to the rest of the data, also introduces a serious flaw since it allows with the foresight of the hypotheses tested, the possibility of 'faking good' - in this case the psi-conducive personality stereotype. (The absence of a motivational distortion scale on form B makes this impossible to check.) Indeed Sargent's and Harley's 16PFs contribute the highest scores among the conducive group on all four of the measures Sargent found to discriminate between conducives and inhibitors. But worse is to come, also among the 'conducives' Sargent somehow includes Drucker

who has not even reported carrying out a single Ganzfeld study (note 4), and a scrutiny of the psi-inhibitory group reveals it to be composed of Palmer, Tart and Rogo (I have already commented on the arbitrary inclusion of Tart and Rogo), Beloff and Millar who conducted a Ganzfeld study using the same subjects, and finally Blackmore who as far as I know has not published her Ganzfeld work. How can one hope to extract any meaning from comparisons based upon this mess purporting to be two homogeneous groups?

In order to recapitulate, it must be stressed that the most serious flaw in this study is the apparent arbitrary way of assigning experimenters to the psi-conducive and inhibitory groups which thereby allows the data to be open to subjective bias in placing experimenters in the groups according to which their 16PF profiles will best confirm the experimenter's hypotheses. By way of illustrating just how much Sargent's data is vulnerable to this kind of manipulation, I re-analysed his scores after performing only two quite justifiable changes: omitting Tart and moving Rogo and Stanford into opposite groups. Both these procedures are justified by the reasoning given earlier. This results in only one of Sargent's four previously significant factors or traits retaining significance - at a marginal level (using one or two tail tests). This is the second order extraversion factor which owes much of its robustness to the hefty input from Sargent's own personality profile! Adhering to the above reasoning further would require that we also omit Drucker with the effect that all the results become nonsignificant. No doubt many other nonsignificant variants are possible.

From these considerations it would seem that without overstating the matter Sargent's research here amounts to little more than a dressed up case study in which results are obtained by fitting the data to the hypotheses. Nevertheless there may be something to be learned. Without such scrutiny it would be all to easy to regard Sargent's findings as further encouragement for the oneupmanship game in parapsychology where nonsuccesses of certain experimenters are explained away by lack of an adequate personality (while the other side explains away successes by sloppy experimentation or fudging). No doubt it is still worth carrying out a more sophisticated analysis than either mine or Sargent's but it is to be hoped that we are coming away from such naive notions as stable unitary personality types underlying the experimenter effect. If there is a psychological explanation to the experimenter effect in parapsychology, then it seems intrinsically more plausible to suppose that it has its basis in interpersonal

TABLE 1
Re-analysis of Sargent's data

Factor	A	F	Q1	Neuroticism
Mean conducive	4.60	5.80	5.86	4.69
Mean inhibitory	2.40	3.20	2.34	6.52
Difference	2.20	2.60	3.52	1.83
SD-difference	1.82	1.81	1.50	1.51
t (df=8)	1.21	1.44	2.34	1.21

factors such as expectancy of success, rapport, and nonverbal communication than in superficial categories such as extraversion and neuroticism.

NOTES

1) It is also worth noting that Saville and Blinkhorn (1976) chose forms C and D for their standardisation of the 16PF amongst university students.

2) The 1977 Rogo study is an example of a study difficult to classify. Two of the 5 subjects produced significant psi-missing and a "partially post-hoc" chi-square was also significant. The Stanford and Neylon study obtained significance only on a secondary analysis for time in Ganzfeld estimates, and the recent Stanford (1979) study failed to obtain any evidence of ESP on a main effects ANOVA. Some secondary effects were obtained but here several of the analyses were questionable being conducted following nonsignificant ANOVA's. In a private communication Sargent (1979) mentions that he regarded psi-inhibitors as experimenters "who do not obtain overall significance in their results". On this basis Stanford and Rogo can be regarded as misclassified.

3) I have recently learned that Sargent and Harley have now had success with individual subjects so this criticism may not apply.

4) I believe Sargent has confused Drucker's sensory deprivation/change in state study which she conducted as a junior experimenter with Honorton, with being a Ganzfeld study. In any case the results cannot be equated with those of a successful Ganzfeld study.

REFERENCES

Ashton, H., Dear, P.A., Harley, T.A., Sargent, C.L. 'A four subject study of psi in the Ganzfeld', J. S.P.R., 1981, 51, 12-21.

Cattell, R., Eber, H., Tatsuoka, M. 'Handbook for the Sixteen Personality Factor Questionnaire', Champaign, Illinois, IPAT Press, 1970.

Parker, A. 'Parapsychologists' personality and psi in relation to the experimenter effect', in Morris, R., Roll, W.G., Morris, J. (eds) 'Research in parapsychology 1976', Metuchen, Scarecrow Press, 1977.

Sargent, C.L. 'A note on personality differences between psi-conducive and psi-inhibitory Ganzfeld experimenters', in Roll, W.G. (ed) 'Research in parapsychology 1979', Metuchen, Scarecrow Press, 1980.

Saville, P, Blinkhorn, S. 'British undergraduate norms to the 16PF', NFER, Windsor, U.K., 1976.

Adrian Parker
Child and Adolescent Psychiatry Clinic
Goteborg
Sweden

A CASE OF THE REINCARNATION TYPE IN INDIA
THE CASE OF RAKESH GAUR

Satwant K. Pasricha
David Read Barker
University of Virginia

The first efforts to investigate in detail the cases of children in India who claim to recall a previous life, began more than 50 years ago (Sahay, 1927; Sunderlal, 1924). It has been only since 1966 that readers in the West have had ready access to case material of this type. In that year Dr. Ian Stevenson published reports of 20 cases suggestive of reincarnation that included reports of seven Indian cases (Stevenson et al., 1974) and detailed reports of 10 additional Indian cases (Stevenson, 1975). More recently, one of us (S.K.P.) has coauthored three brief reports of Indian cases with him (Pasricha and Stevenson, 1977).

Stevenson has indicated that the cases of the reincarnation type are not unique in their occurrence but can also found by other investigators. He wrote, "I can put my finger on several countries on a map and positively guarantee that if an investigator goes to the place indicated he will find cases similar to those in this book within 50 kilometers, or even 10, of the place I point out" (Stevenson, 1975, p. 3).

Our purpose in presenting this case is twofold: first. it is similar to cases from India that Stevenson has interpreted as suggestive of reincarnation, and second, it is the most thoroughly investigated case of the type that we have studied together without Stevenson's

participation.

We do not regard the investigation of this case as completed. The testimony of the informants showed some variations, and later we each interpreted the significance differently. We have not yet discussed these discrepancies with informants. If it is feasible to do this in the future, the results may justify a supplementary report of the case. it is, however, sufficiently investigated at present to warrant this report.

Methods

In his present book reporting 10 Indian cases, Stevenson (1975, pp. 8-44) has described in detail the methods he developed for the investigation of cases of the reincarnation type.

Essentially, the methods employ repeated, independent interviews with as many firsthand informants as possible, first in the village or town of the subject and then in the place where he (or she) claims to have lived in his (or her) previous life. In addition, all available documents relevant to the case are examined.

We followed Stevenson's methods as closely as we could. Our major deviation from his usual method was that we began the investigations of the case by interviewing members of the family of the person who was said to have reincarnated, rather than with the family of the subject. This change was made because it fitted more conveniently into our travel schedule.

In interviewing we allowed each informant to first give information in his (or her) own way without much interruption. Direct or, on rare occasions, leading questions were later asked to fill any gaps in information required for the overall evaluation of the case.

All the interviews were conducted in the Hindi language. Although D.R.B. knew enough Hindi to follow the gist of the conversation, the literal translation was essential for precision and proper understanding of the case, so S.K.P. translated the conversation for D.R.B. during the interviews. S.K.P. also took verbatim notes in Hindi while D.R.B. took notes in English. In addition, some of the interviews were tape recorded and the tapes later transcribed and compared with the written notes made during the interviews.

We investigated this case together for a total of seven days between November, 1976 and November, 1978. S.K.P. investigated it further by making another visit to Tonk for two more days in March, 1979.

CASE REPORT

Summary of the case and its investigation

Rakesh Gaur was born on March 15, 1969 in the town of Fateh Nagar in Rajasthan (40 kilometers southeast of Kankroli). He is the fifth child, and third son, of six children of S.N. Gaur and his wife, Shanta Devi.

S.N. Gaur, a Brahmin by caste, is an upwardly mobile, middle level employee of the Indian Post Office. He was transferred several times to different posts in southern Rajasthan during the late 1960s and the 1970s. Before moving to Kankroli (District Udaipur, Rajasthan) in 1975-76 , the Gaur family lived in Fateh Nagar and Bhinder (70-80 kilometers southeast of Kankroli). S.N. Gaur was the Sub-Postmaster of Kankroli and resided there with his family during our investigation of the case in 1976 and 1978.

In May, 1974, when Rakesh was little more than five years old, he began to tell his parents that in his previous life he had been a member of the carpenter caste who had lived in the Chhippa neighborhood of Tonk (also in Rajasthan, about 225 kilometers northwest of Kankroli), that he had been married to a woman named Keshar, and that he had been electrocuted. Three informants said that Rakesh mentioned the name of Bithal Das before the two families met, although three other informants said that he had not done so in their presence. Rakesh showed an interest in carpentry and repeatedly asked to be taken to Tonk. Family friends urged S.N. Gaur to take his son to that city, one day's travel, but he made no definite move to do so.

In June or July, 1976, a bus driver from Tonk named Chhittar Mal (alias Chhittarji) stopped in Kankroli. Chhittarji said that Rakesh recognized him spontaneously and narrated his previous life memories with enough specificity so that Chhittarji returned to Tonk and told the family of Bithal Das that he had been reborn in Kankroli. As a result of this fortuitous encounter, S.N. Gaur wrote to the Tonk

Electricity Board to enquire if any employee belonging to the carpenter caste had been electrocuted. He also wrote to Chhittarji, but he received no reply to either of his letters.

On Rakesh's repeated request and the urging of several friends, S.N. Gaur finally decided to go to Tonk. In October, 1976, when he had leave from his work, S. N. Gaur took Rakesh there in order to confirm his son's assertion that he had lived a previous life in Tonk.

As their bus was entering Tonk late in the afternoon, Rakesh pointed to an electric pole and said that he had died while repairing it. Since S.N. Gaur was a post office employee and knew no one in Tonk, he went to the post office to enlist the help of local post office personnel in tracing the family of the person about whom Rakesh had been talking. Several post office employees set out on foot with Rakesh and his father to try to locate the house of the person in question. A large crowd gradually formed, and one elderly person in the crowd recalled that a carpenter named Bithal Das had been electrocuted in 1955.

Rakesh and the crowd that had gathered around him returned to the post office. They were within 200 meters of Bithal Das's house but did not visit it. Bithal Das's widow, Radha, was sent for. On his way back to the post office, Rakesh saw and recognized (but did not meet) Bithal Das's son, Bhanwar Lal, who was some distance off and leaving for Jaipur; at the post office Rakesh met Radha, who cried most of the time they were together. Although it was nearly midnight, the crowd continued to grow and Rakesh became nervous. Believing that he had confirmed his son's statements, S.N. Gaur returned with Rakesh to Kankroli on a late night bus. They had been in Tonk about six hours.

Two days later Bhanwar Lal visited Rakesh at his house in Kankroli to study the case for himself. Later that month Rakesh was also visited by Bithal Das's elder brother, daughter, and son-in-law.

We learned of this case on November 12, 1976 during our investigation of other cases of the reincarnation type. An associate in Jhalawar (eastern Rajasthan) gave us an article, from the Hindi language newspaper 'Dainik Navjyoti', October 17, 1976, which described in about 400 words Rakesh's visit to Tonk. Because this event had occurred less than one month earlier, we began our investigation of the case as quickly as we could. On November 15, 1976, we visited Tonk (since it came first on our way from Jhalawar to

Jaipur, where we had previously planned to investigate other cases). In Tonk we interviewed three persons of Bithal Das's family, namely his son Bhanwar Lal, his widow, Radha, and Radha's younger sister, Ladha. On November 18, 1976, we arrived in Kankroli and interviewed Rakesh and his parents (note 1). Subsequently, in February, 1978, we returned to both these places and interviewed all the previous informants (except Ladha) as well as some new ones who were familiar with the case. In November, 1978, we returned to Tonk, mainly to interview Chhittarji, who was a key witness in the case (we had not met him on previous occasions); but also to meet employees of the Post Office and Electricity Board who had been present during Rakesh's first visit to Tonk or who had knowledge of the postcard that S.N. Gaur had written to the Electricity Board. In addition, we went to Bundi (110 kilometers east of Tonk) to meet with Bithal Das's elder brother, Ram Bilas. In March, 1979, one of us (S.K.P.) returned to Tonk once again in order to verify some of the items that we had listed previously and to interview a few new informants.

In all, we conducted 36 interviews with 24 informants. Two persons were interviewed four times, six were interviewed twice, and 16 were interviewed once. In addition to eleven members and five friends of the two families concerned, we interviewed eight informants who were not directly related to either Rakesh's or Bithal Das's family but who were witnesses to the statements or recognitions that Rakesh made during his first visit to Tonk.

Life and death of Bithal Das

Bithal Das was born in Tonk in about 1922, one of five children in his family. From 1937 to 1942, when he was between the ages of 15 and 20 years, he resided in Bombay with Ram Bilas and worked as a carpenter and chauffeur. In 1942 he returned to Tonk, where he opened a small carpentry workshop with Ram Bilas in the Kali Paltan ward. This venture lasted five years, until 1947, when Bithal Das moved to the small city of Bundi and opened up another carpentry workshop there. In 1950 Bithal Das returned to Tonk complaining of 'weakness', saying that his doctor had advised him to rest.

Bithal Das was married while still a teenager to Radha, a local girl who had also grown up in a 'sutar' (carpenter) family. They had five children, of whom two sons and two daughters lived beyond infancy. All

the children were born in Tonk, the eldest, Bhanwar Lal, having been born in 1942, shortly after Bithal Das returned to Tonk from Bombay.

There have never been many carpenters in Tonk in recent years. At the time of Bithal Das's birth in about 1922, there were about 40 to 50 carpenter families, but their numbers have declined slightly, and there were only 30 to 35 'sutar' families in 1978. Bithal Das's was the only 'sutar' family residing in the Kali Paltan ward during the 1970s.

Bithal Das was described by a friend as, "very calm, very religious, and a strong supporter of the Congress Party. He was a good mechanic but a poor person". According to a relative of Bithal Das, "as a carpenter, his skills were good among Tonk people but economically he was not so well off." He was interested in electrical work (his elder brother called it 'his hobby') and used to do it privately. He had done some wiring at his sister's house around the time of her wedding. (The family had gone to attend her wedding at the time of Bithal Das's death.) He was never employed as an electrician, however, and never worked for the Electricity Department.

During the rainy season, on August 15, 1955, Bithal Das was alone at home. Some water had gathered in the drain and courtyard of the house. He was clearing the blockage with an iron rod, which accidently touched a live electric wire. Bithal Das received a shock and died instantly. The family's Brahmin priest was called to officiate at his cremation, which was attended by the male members of his community. He was about 34 years old when he died.

Geographical factors and possible connections between the two families concerned

Kankroli is a large town in the north Indian state of Rajasthan, located on National Highway 1, 60 kilometers north of the city of Udaipur and 300 kilometers southwest of Jaipur, the state capital. Tonk is a small city with a population of 70,000 located 80 kilometers south of Jaipur. The straight-line distance between Kankroli and Tonk is 225 kilometers. There is no regular exchange of goods or people between Kankroli and Tonk.

No informant indicated that there had been any kind of contact

whatsoever between the family of Bithal Das and the family of Rakesh Gaur before Rakesh began to speak about a previous life. S.N. Gaur, Rakesh's father, had heard of Tonk before his son spoke of a previous life there, but he had never been there, did not have any relatives there, and did not even know exactly where it was located.

EXTRACTS OF TESTIMONY

In order to present this case as authentically as possible, we shall next give some extracts of the testimony we obtained. Nearly all of the extracts included in this section are exactly as they were recorded in our notes. In a few instances, minor editorial changes have been made to express the statements in more idiomatic English without changing the meaning of any statement. We present the testimony in chronological order of the development of the case, which is not necessarily the order in which we recorded the different statements.

The first statements by Rakesh

In November, 1976, S.N. Gaur described the first statements made by Rakesh that were later understood to relate to a previous life:

> About two and one-half years ago, when we lived in Bhinder, Rakesh sometimes used to say, "I belong to Tonk. I am a 'sutar' (carpenter)". We did not pay any attention.
>
> During the rainy season, Rakesh used to wake up in the night sometimes at 1 or 2 a.m.. We used to ask him what the matter was. He did not tell us in the beginning. One day we insisted, and he said, "It is raining so heavily over here, and it is causing so much damage, so many houses are collapsing. What will become of my children and my wife? My house was made of unbaked brick." Then he started giving other details, including the name of his neighborhood, Chhippa, and his own name, Bithal Das. He said, "I was electrocuted. I had two sons. I had a wife. Her name was Keshar. There is a well near my house called 'bara kunwa' (meaning 'big well'). We used to get water from there".

> Then I told him that all of us would go to visit them. He said, "Not all of us. We were very poor and used to eat only pulses. We never ate any vegetables. If all of us go, what would they give us to eat? We shall stay there only for one night. Papa, only two of us, you and I, should go."

Adjoining the post office and staff housing in Kankroli is a carpentry shop owned by Mr. N.L. Sharma. In February, 1978, he told us:

> Before he first went to Tonk, Rakesh used to talk about his previous life quite often. He used to come and play here, and he took some interest in carpentry. Sometimes he used to tell the carpenters about the tools, and he used to give them instructions (such as) "use this tool like this and that tool should be used like that." Then he said, "In my previous life I was a carpenter. I used to live in Chhippa neighborhood, in Tonk." I asked him how he died, and he said, "I was working on an electrical pole and fell from it. I was wiring the pole and had a shock and died."
>
> I was quite surprised when I heard these things from him, so I told his father that he should take Rakesh to Tonk. I had heard of (reincarnation) cases sometimes, and I had read about them in the papers, but I had had no personal contact with a case. I was quite surprised that such a small child should tell about his previous life.

The recognition of Chhittarji

Despite the insistence of his son and the urging of several friends that he take Rakesh to Tonk, S.N. Gaur might not have taken any action had it not been for a fortuitous meeting between Rakesh and a bus driver from Tonk called Chhittarji. In November, 1978, Chhittarji described this meeting, which probably occurred in June or July of 1976:

> My bus and three or four others went to Ram Deora. On the way to Nath Dwara, with pilgrims from Tonk, I stopped in Kankroli and was standing by a shop when that boy recognized me. All the buses had 'Tonk' written on them, but the boy

> did not see the buses. He said to his father, "This is Chhittar." I didn't pay any attention. Then that child called me and said, "I have two children there (in Tonk). One is Bhanwar and the other is Babu." Then that child asked me to go to his house with him and said, "You go and tell my children. Haven't you recognized me? I am Bithal Das." I didn't pay any attention. I thought he was a child and was just talking like that. Then I came back to Tonk and gave the message to his (Bithal Das's) sons that there is a child in Kankroli who says, "I am Bithal Das." They didn't pay any heed to it. Then I received a letter (actually a postcard which S.N. Gaur later wrote him; see below) and gave the letter to Bithal Das's sons.

The wife and oldest son of Bithal Das narrated to us a different version of this recognition. But their account is second hand and should be viewed in that perspective. According to them, Chhittarji has told them that some shopkeepers in Kankroli, seeing that his bus was from Tonk, asked, "There is a child here who says, "I am from Tonk and was electrocuted. I had two children." Is there any such person in Tonk?" Rakesh's father then called Chhittarji to the post office and asked Rakesh, "Do you recognize him?" Rakesh replied, "Yes, I taught him how to drive."

Although these accounts of Rakesh's meeting with Chhittarji are different, they are not necessarily incompatible. Rakesh's father might also have called Chhittarji to meet Rakesh at the post office. Be that as it may, the fact remains that Bithal Das had, according to Chhittarji, taught him both carpentry and driving. Chhittarji, who was about six years younger than Bithal Das, had lived with him and was present in the neighborhood at the time of his death.

S.N. Gaur wrote to Chhittarji requesting him to verify what his son had said. Not content with this, he made a second effort to confirm his son's statements through correspondence. Because Rakesh had talked about being electrocuted while wiring a pole, S.N. Gaur wrote a postcard to the Rajasthan State Electricity Board in Tonk asking if a carpenter had been electrocuted. The postcard was turned over to two clerks named Phool Chand and Bhol Chand who searched the records but found no employee of the Board who had been electrocuted. Unfortunately, the postcard was thrown away. Phool Chand insisted, when we interviewed him in November, 1978, that Bithal Das's name was given in the postcard and that employees of the Electricity Board knew

of his death but that since he had not been an employee of the Board, there was no official record of his electrocution in the departmental files. Bhol Chand recalled the postcard and said that it included the name of the person being enquired about but that he could not remember the name.

Rakesh's first visit to Tonk

Since S.N. Gaur received no reply to either of his communications, he decided to take Rakesh to Tonk in order to confirm his son's assertion that in a previous life he had lived there. This visit, which took place on October 16-17, 1976, lasted only about six hours, from 7:30 p.m. to 1:30 a.m..

In an effort to reconstruct the precise course of events during this brief visit, we interviewed six informants at length: Rakesh's father, the widow and son of Bithal Das, two senior postal employees (B.K. Bhatnagar and H.P. Sharma) who were in Tonk at the time of the visit, and an accountant (Mahi Lal) in the post office. These three postal employees accompanied Rakesh during a walk through the city. All the informants, with the exception of Mahi Lal, substantially agree in their account of the visit. Therefore, we give the testimony of S.N. Gaur, who was with Rakesh throughout the visit, and of Mahi Lal, who gave a different description of the events. We also include a short extract of the testimony of H.P. Sharma, assistant postmaster, who said that he had witnessed the events during this visit that led to the identification of Bithal Das. In November 1976, one month after the visit, S.N. Gaur described it as follows:

> When the bus entered Tonk, there was an electric pole by the road. Rakesh immediately said, "Papa, this was the pole were I met with an accident."
>
> So I got some confirmation that I would surely know the truth here.
>
> I thought, let's go to the post office. We reached there about 7:30 p.m. and told the postal workers that Rakesh says so-and-so. We took a postman with us whom we thought would help us in locating the house. Well, Rakesh used to say that

he was 35 when he died. I asked him, "How long ago?" He answered, "I don't know, but I was 35 when I died." We could not locate the house. Rakesh said, "I can reach my house", and started walking ahead of us. After a while something went wrong with the lights. So I thought that since we were not in a hurry, we would try in the morning.

A large crowd had gathered by then. Someone in the crowd said, "I know of a family of 'sutars' (carpenters) where a man died at the age of 35. If you wish I can call the family." Meanwhile someone from the crowd informed the wife of Bithal Das. Two ladies were coming together, and Rakesh immediately pointed at his wife among the other ladies, saying, "She is my wife." The crowd had become so thick that it was impossible to talk there. So we told the ladies to come to the post office. But when we reached the post office, it was even more crowded. Rakesh became nervous and I got worried that he would fall ill. So we dispersed the crowd, telling them that we would talk in the morning. Whatever the post office staff wanted to know from Rakesh, they asked. I could not leave him alone.

Then we went to have dinner at the Inspector's home. When we returned to the post office, it was 11:30 p.m.. The crowd was still gathering and it became a difficult situation. All of us in the post office agreed that we had found (verified) everything that Rakesh used to tell about. What more did we want to know? I wanted only to confirm that much. So we came back (to Kankroli).

Mahi Lal, an accountant at the post office, gave, in February, 1978, the following account of the events of the visit:

In the post office here, somebody asked, "Would you be able to recognize the house?"

The boy said, "I was a lineman in the electric power house. I died as a result of electric shock." The lineman's caste was 'khati' (same as 'sutar'). The post office staff was asked whether there was such a death in that area. One of the postmen said, "Yes, there was such a death 20 to 25

years ago."

Around 6:30 p.m., some staff members and the boy started towards the house. First he took us into a wrong lane. When we reached the right house, we were told that they had moved away. We kept asking along the way. When he was taken to the place where that man (Bithal Das) had died, even then he couldn't recognize the place. Someone else told us that in Kafala neighborhood there was a 'khati' who died like that. When you come from Old Tonk on the Jaipur Road, there is an electric pole. Seeing that pole, he said, "I died when I was repairing that pole." His father said that at home the boy says he died of electrocution. After that, he was brought to Kafala neighborhood. We didn't reach the house, and we brought the child back (to the post office). We sent word that the wife of Bithal Das should come to the post office. The boy was asked, "Do you recognize her." He didn't say anything. He just turned his face away from her. H.P. Bhatnagar asked, "Is she your wife?" Then the boy said, "Yes." After that she left.

The child became nervous. The postmaster advised that they not stay in Tonk because the crowd was growing so large. They left on the night bus at about 2 a.m..

The boy never said what his name was, only that he was a carpenter with two sons and had died at an electric pole.

The contradiction was that the boy said that he had died at a pole, but the death had occurred at home. He recognized the pole as well as the house - that is the contradiction. He said he had two children, and that was correct (note 2).

In February, 1978, H.P. Sharma, assistant postmaster, gave his account as follows:

Eight or nine of the post office staff and several other passers-by accompanied the boy. He was guiding us, and we followed him. It was dark by then, but the boy continued to walk in the streets of Old Tonk. Suddenly the boy stopped and told us that he used to live in the proximity of that

street, but it was not possible for him to exactly locate the house. The neighbors were contacted, but none could recollect a person who corresponded with the boy's statements. Then one elderly carpenter said, "Certainly, there was a person of that name who used to work at the Electricity Board who died of electric shock." The family of the deceased carpenter had moved to New Tonk, and we proceeded, following the boy, to New Tonk and reached the house of the carpenter and met his wife, named Keshar (note 3), and her two sons.

The first visit by Bithal Das's family

Bhanwar Lal, Bithal Das's oldest son, was the first in his family both to hear of and see Rakesh, and he was the first to visit Kankroli, two days after Rakesh returned from Tonk. A month later, in November, 1976, he described his visit:

When I first reached Kankroli, the child (Rakesh) gave the background entirely wrong (note 4). He said that his name was Arun, that he had lived in Tonk, had been electrocuted, and had two sons. He didn't mention daughters. I asked him two questions.

I asked, "How did you die?" He said, "When I died there was no one home. I kept 1500 Rupees in the house." That was correct.

My second question was about the house. I asked, "How was your house?" I had in mind that he would say 'kachcha' (made of unbaked brick) or 'pukka' (built of baked brick), but he gave the whole situation. (That is, Rakesh went on to describe details of the house spontaneously.) He said, "Outside my house is an electric pole. After that there is a well, then a gate, and then a market." Now the well has been filled. He said, "There are four rooms in the house." This is correct. He said, "I had one room to myself which was being painted green. The room had two doors. There were two

> windows with four swinging panes. In the center table there was kept 1500 Rupees." All this was correct.
>
> The third thing he said spontaneously. He said, "When I lived in Bombay I used to get green 'sarees' (dresses) for your mother. You should confirm this from your mother. I used to eat 'urad' (a kind of pulse).

This visit, although it lasted only about four hours, was sufficient to convince Bhanwar Lal that Rakesh was the reincarnation of his father, and he returned to Tonk to tell his family about this.

Later behavior and development of Rakesh

In the two years following his first visit to Tonk, the families of Rakesh and Bithal Das established close ceremonial ties through an exchange of visits. The relationship between the families grew in three stages. During the month following Rakesh's trip to Tonk, he was visited in Kankroli by Bhanwar Lal, the eldest son of Bithal Das, by the elder brother of Bhital Das, and by a daughter and son-in-law of Bithal Das. These visits were mainly made out of curiosity to know the facts, all lasted less than one day, and none of these people stayed with the Gaur family overnight. Five months later, during April-May, 1977, Bhanwar Lal escorted Rakesh and his father to the wedding in Bundi of Bhital Das's granddaughter. All of these meetings were characterized by growing conviction on the part of the participants that Bhital Das had been reborn as Rakesh.

The case quickly became a minor sensation in the Indian press. Extraordinary powers of clairvoyance and precognition began to be described to Rakesh, and he was treated as a special child, a development that occasionally occurs in cases of this type (note 5).

DISCUSSION

As we have reviewed the data of this case and discussed it between ourselves and with other persons, we have felt that although we agree as to the main items of testimony to be evaluated, we disagree

concerning the weight to be given to some of these items. This has led us to prefer different interpretations of the case, and we differ on these so much that we think it best to write parts of this discussion separately.

Before coming to our different interpretations of the case we will summarize certain features of the case about which we are agreed.

We both believe that the informants are all fundamentally honest people who have not tried to deceive us; whatever mistakes they have made we attribute to faulty memories and unconscious biases in observing and reporting the events they witnessed. Further, we also believe that the families of Rakesh and Bithal Das had had no contacts, direct or indirect, before Rakesh met Chhittarji. Rakesh had never been to Tonk before meeting Chhittarji.

Interpretation by D.R.B.

The major constraint on the interpretation of this case is that no written record was made of Rakesh's statements about a previous life; the only documents written before his first visit to Tonk, the postcards from his father, were thrown away. An authentic reconstruction of the case therefore depends entirely on the capability of our informants to observe, remember, and report the events as they actually occurred and on our ability to elicit and accurately record their memories.

The informants varied widely in education, intelligence, capacity for verbal expression, and emotional involvement in the case. Without exception, the eleven members of the families of Rakesh and Bithal Das whom we interviewed had a deep emotional investment in their understanding of Rakesh as the reincarnation of Bithal Das. Of the remaining informants, all but one appeared to believe strongly that this was a case of recall of a previous life. The members of Bithal Das's family venerate Rakesh. Rakesh's father, the most important informant, appeared to have an unstable attitude toward his son that include awe, indifference, annoyance at the cost and inconvenience of the numerous inquiries the case generated, and pride. We noted that he seemed to have a tendency to exaggerate slightly.

The most striking feature of the testimony in this case is that

there are a great number of discrepancies and contradictions in informants' descriptions of events leading up to the identification of Bithal Das as the person whose life Rakesh was describing, but there are almost no discrepancies in the testimony concerning events following Rakesh's first visit to Tonk. Agreement among informants is relatively low regarding facts of the early stages of the case, but almost unanimous concerning later events. This feature of the testimony cannot be attributed to the interval between the time of the events and their report. It is instead a reflection, in my opinion, of the uncertainties, misunderstandings, and confusion that surrounded Rakesh's early statements and visit to Tonk.

The rapid transition from highly discrepant to nearly unanimous testimony parallels the degree of accuracy of the statements about his previous life and the facts of Bithal Das's life. During and after the exchange of visits following the first visit to Tonk, however, no one reported that Rakesh had made incorrect statements; every statement and recognition, without exception, was said to have been correct.

With these general observations on the characteristics of the testimony as a background, I will first review those statements that are unanimously attributed to Rakesh and that correspond to the facts of Bithal Das's life. Then I will summarize those features of the case that are disputed by one or more informants and/or do not correspond to Bithal Das. Finally, I will touch on the statements and recognitions made by Rakesh after he gained access to normal knowledge of Bithal Das's life.

All the informants agreed that Rakesh said three things corresponding to Bithal Das's life: (a) He had belonged to the carpenter caste, (b) he had lived in Tonk, and (c) he had been electrocuted. Further, it is clear that Rakesh, during his first meeting with Chhittarji, said enough about his previous life memories for Chhittarji to tell Bithal Das's family that Bithal Das had been reborn in Krankoli. Chhittarji said that Rakesh recognized him spontaneously, saying, "I know this man, I taught him driving."

The most serious point of disagreement among the informants is whether Rakesh actually said he had been named Bithal Das. Three major informants, including Rakesh father, insisted that he had. Three other major informants, including Bithal Das's son, said he had not.

The events of the first visit to Tonk suggest to me that Rakesh had

not given the name Bithal Das before going to Tonk. He and his father spent about six hours searching Tonk for his house in the previous life. It seems to me extremely unlikely that the boy would have been asked to do this if the name Bithal Das had been attached to his other statements from the moment of their arrival; the postmaster would have simply sent someone the 500 meters to Bithal Das's house to tell a member of his family to come immediately. This is what eventually happened, but only hours later, after an elderly person in the large crowd thought of Bithal Das, whose wife was then brought to the post office.

The testimony in this case indicates clearly that Rakesh made some statements and 'recognitions' that did not correspond to the life of Bithal Das. Rakesh said that his wife in his previous life was named Keshar; Bithal Das's wife was named Radha, alias Mallo. Rakesh said he had lived in Chhippa neighborhood of Old Tonk; Bithal Das lived in the Kali Paltan neighborhood of New Tonk. Rakesh said that his house was built of unbaked brick and that it was not electrified; Bithal Das lived in a house of baked brick that was served by electricity. It was at his own home, in fact, that he had been electrocuted. Rakesh 'recognized' as the site of his death an electrical pole that is located about 2 kilometers from Bithal Das's home, and he 'recognized' his own home in a neighborhood entirely different from Bithal Das's.

In the light of these discrepancies, the testimony of Bhanwar Lal, Bithal Das's son, is particularly important. According to him, Rakesh described his house in the previous life accurately and in great detail when Bhanwar Lal visited him in Kankroli two days after Rakesh's first visit to Tonk.

The disparities between the previous life recalled by Rakesh and the life of Bithal Das go beyond the specific differences noted here. Rakesh seems to have been talking about an extremely poor individual before his visit to Tonk. He went as far as asking his mother and siblings not to accompany him and his father to Tonk because, "we were very poor and used to eat only pulses. We never ate any vegetables. If all of us go, what would they give us to eat?" This expression of poverty was consistent with his earlier fears that his unbaked brick house would collapse in the monsoon rains. Bithal Das, however, had managed to pull himself out of this level of poverty, and in the last years of his life he was a lower middle-class urban householder with a brick house and a significant hoard of precious metal.

I interpret the case of Rakesh Gaur largely as a product of Indian social psychology and the widespread belief in the possibility of remembering a previous life. As I reconstruct this case, Rakesh as a young child repeatedly made a few general statements about a 'previous life'. He was taken to Tonk in order to 'confirm' these statements by leading his father to the family of his previous life. He could not do this because many of his statements and recognitions did not correspond to anyone known to have existed. After two or three hours of searching, a large crowd gathered. A person in the crowd thought of Bithal Das, a man who matched enough of Rakesh's statements to satisfy nearly everyone that Rakesh had been talking about this man from the beginning. From that moment Rakesh was provided with abundant cues for appropriate behavior, and Rakesh, an unusually intelligent and observant child, quickly learned to become Bithal Das. Equally important, everyone who was personally involved in the case became convinced that Rakesh actually was Bithal Das, and they treated him accordingly. In consequence, all the people directly involved with Bithal Das and Rakesh rapidly created a reality that is deeply rooted in Indian psychology. Rakesh became the reincarnation of Bithal Das because he identified himself fully with Bithal Das's life and because other people supported him and rewarded him in this identification.

Although 'orthodox' social psychology provides, in my opinion, an adequate framework for understanding the characteristics of the testimony noted above and the process by which Rakesh came to be identified as the reincarnation of Bithal Das, there remains at least one statement attributed to Rakesh that does not seem to have a normal explanation. This is his recognition of the bus driver Chhittarji. Although we cannot reconstruct their meeting with certainty, it is clear that Rakesh said enough to Chhittarji about his former life for Chhittarji to have told Bithal Das's family that Bithal Das had been reborn in Kankroli. Three members of Bithal Das's family confirmed this. Furthermore, Chhittarji's behavior resulted from only a short meeting (perhaps ten minutes) with Rakesh, during which time Rakesh somehow told him enough to identify Bithal Das specifically.

Stevenson's guarantee of the existence of cases of the reincarnation type in certain parts of India is partially correct. The characteristics of Rakesh Gaur's case are very similar to those in cases he reports as probable instances of previous life memories. Those readers familiar with Stevenson's detailed case reports will almost certainly feel that Rakesh Gaur is another subject of a 'case of the reincarnation type', at least from the viewpoint of social

anthropology. Although I do not interpret this particular case as evidence that a previous life can be recalled, there is in the case at least one instance of what might be paranormal knowledge on Rakesh's part.

Retrospective reconstructions of complex events, exemplified in this case report, will be transformed and strengthened immensely through prospective methods. I hope that as the importance of these cases is perceived by scientists, qualified investigations can begin before the two families meet. Had we been able to accompany Rakesh during his first visit to Tonk, we might have acquired stronger evidence that reincarnation actually occurs.

Interpretation by S.K.P.

My colleague and co-author quite rightly draws attention to Rakesh's mistakes and to inconsistencies in the testimony of informants for this case. Although these seem to me slightly more numerous, on the average, than those occurring in other cases that I have investigated, I would not characterize them as 'many' or 'a great number'. The discrepancies become important not by their number, but by the persons and circumstances connected with their occurrence. If informant A reports that the subject made a particular statement to him, but informant B says that the subject did not say the same thing in his presence, their testimonies are not discrepant unles both were referring to the same occasion when both were present. A subject may make one remark to one person, but not repeat this in the presence of another person on some other occasion. It seems elementary also that we must evaluate the respective credibility of two informants who give discrepant testimony, even when they both are present on the same occasion. If, for example, it is evident that one informant was paying little attention to what went on, his testimony should not descredit the statements of another who shows himself more deserving of our confidence. I also think it important to distinguish variations in unimportant details from true contradictions in different accounts. With these principles in mind, I will itemize all the statements credited to Rakesh that bear on the question of whether knowledge he had before going to Tonk led to the identification of Bithal Das as the person about whom he was talking. In considering this question, we must ask ourselves whether the statements attributed to Rakesh apply specifically to Bithal Das rather than to someone else who happened to

die of electrocution in Tonk. In what follows, I deliberately ignore Rakesh's recognitions (except for his mistaken recognition of an electric pole) since these are particularly liable to error on the part of informants. In any case, recognitions played little part in the question of whether Rakesh had substantial knowledge about Bithal Das before he first went to Tonk.

The identification of a person corresponding to Rakesh's statements

According to both S.N. Gaur and Chhittarji, Rakesh told them at different times that his name in the previous life was Bithal Das. I will call this item 1, and number sequentially the other items to be considered later. According to Phool Chand, an employee of the Electricity Board of Tonk, S.N. Gaur had mentioned this name in the postcard of inquiry he sent to the Board; another employee of the Electricity Board, Bhol Chand, also said that the postcard included the name of the person being enquired about, but he could not remember the name.

S.N. Gaur, as well as several other witnesses, testified to various statements Rakesh made before his first visit to Tonk. (They did not, however, all testify to the same items.) These were: that (in the previous life) he (2) had been a carpenter of Tonk; (3) had been married with two sons; and (4) had died of electrocution (5) at the age of 35. Rakesh also referred (6) to a 'bara kunwa' (big well), which he described as 'near' his house. To Chhittarji he said that (7) he (Bithal Das) had taught him (Chhittarji) driving, and he gave the names of his sons as (8) Bhanwar Lal and (9) Babu. He also said that (10) he had hidden 1500 Rupees at his house. All these statements were found to correspond to facts in the life of Bithal Das with the exception of the reference to the 'bara kunwa' as 'near'. (I shall return to this detail later.)

Rakesh also made (before the families met) two important statements that were incorrect for Bithal Das. He said his wife (of the previous life) was called (11) Keshar and (12) that he had been putting wire on a pole when he was electrocuted. The second of these mistakes led S.N. Gaur to assume that the person about whom Rakesh was talking probably was an employee of the Electricity Board engaged in some official work as an electrician when he was electrocuted. Rakesh made a further mistake when he said that (in the previous life) (13) he had lived in

the Chhippa Mohalla.

Rakesh also made some other statements about Bithal Das's economic conditions and house to his father. He had said that (14) he was very poor; that (15) his family ate only pulses (they never ate any vegetables); and that (16) the house was 'kachcha', meaning of inferior, clay construction. The latter statement harmonized with his repeated expressions of concern about what rain might do to his house and children. He said that (17) the house was near a 'nala' (channel of water) and implied that the area could be flooded during heavy rains. In fact, at the time of his death, Bithal Das lived in a house made of 'pukka' (brick construction). His old house, however, across the street from the new one, was partly of 'kachcha' construction, and Bithal Das and his family continued to occupy a portion of this house. This living arrangement also indicates Bithal Das's fluctuating economic conditions. He was described by a friend and by a relative as 'poor' and 'not so well off'; his wife also told me that although they cooked both pulses and vegetables, they almost never ate both for a single meal. The family cooked their meals in the kachcha house, and, perhaps more important, his children played in the compound of that house. These houses are located in an area that is subject to flooding during heavy rains; in most parts of Tonk, however, flooding is somewhat unusual. Bhanwar Lal told me that even the bazar (located about 200 meters away) is not flooded during heavy rains. Rakesh therefore gave an accurate account of the situation of Bithal Das's house and its vulnerability during heavy rains and their attendant floods.

I think it also worth noting that Bithal Das's possession of a brick house is no indication that he was prosperous. Moreover, the testimony of his wife and a friend (mentioned above) indicates that he was indeed 'poor', as Rakesh said. This poverty may have resulted from his following the doctor's order to 'rest' during the last five years of his life. Yet Bithal Das was obviously not destitute or he would not have had 1500 Rupees in a table of his house when he died.

As mentioned above, Rakesh alluded to (6) a large well ('bara kunwa'), which he said was 'near' his house. There was a well (item 18 below) within 50 meters of Bithal Das's house, and his family had formerly used it. (It had been covered up a few years before Rakesh's first visit to Tonk.) This nearby well gave hard water, however, and so the family had to draw water from the 'bara kunwa'. The latter well was said to be located between 1 and 2 kilometers from Bithal Das's

house. Some may consider this 'near' the house; perhaps those carrying the water, principally Bithal Das's wife, would not.

From the informants' accounts it is evident that the two families had not met before the development of the case. We also know that they met only cursorily at the time of Rakesh's first visit to Tonk. Nor did Rakesh actually reach Bithal Das's house in Tonk, although he probably came within 200 meters of it when he approached the great gate to the bazar beyond which Bithal Das's house is located. On this occasion, members of the two families were within sight of each other, and, according to both S.N. Gaur and Bhanwar Lal, Rakesh recognized Bithal Das's wife and Bhanwar Lal. This meeting, however, occurred in the late evening, and, although there had been some understanding that they would meet again the next day, S.N. Gaur decided that he had verified enough of what Rakesh had said, and he returned to Kankroli without any further meeting with Bithal Das's family. Therefore, although Rakesh might have learned some details about Bithal Das's house and situation from members of the crowd in Tonk, he did not learn any information on this occasion directly from Bithal Das's family itself. Immediately after their brief encounter with Bithal Das's wife and son, Rakesh and his father withdrew to the post office, where Rakesh was further interrogated by the postal employees. Since none of them had earlier been able to identify the person or family to whom Rakesh's statements referred, they were probably not in a position to furnish Rakesh with details about the location and inside features of Bithal Das's house. Additional evidence for this conclusion comes from the journey Bhanwar Lal took two days later to meet Rakesh in Kankroli. He would not have made this journey if the two families had already exchanged information, and when he reached Rakesh at Kankroli he questioned him about various details concerning the life and affairs of Bithal Das. Bhanwar Lal must have been certain that neither he nor his mother had given information about these details to Rakesh during their brief meetings on the night when Rakesh had first come to Tonk. He must further have assumed that no one else (say, in the crowd at Tonk or later at the post office) had furnished Rakesh with details about Bithal Das's house. (Only someone with a knowledge of the inside of the house could have done this.) Bhanwar Lal said that when he asked Rakesh about the house he had expected Rakesh to make only a simple statement about its structural quality; but Rakesh spontaneously went on to give correct details of the house's interior. He gave the following items: (18) Going from the house toward the market one came to a well; (19) between the well and the market there was a gate; (20) an electric pole stood outside the

house; (21) the house had four rooms; (22) he (Bithal Das) slept in a room that was painted green; (23) this room had two doors; (24) this room also had two windows, each with four swinging panes; and in a table of this room were 1500 Rupees (10). With one exception, I verified all these items myself from an examination of the house and its location in relation to the nearby well and market. (I did not actually see the 1500 Rupees, but I can nevertheless think of no reason why this item should not be credited, especially since Bithal Das's son, Bhanwar Lal, acknowledged as correct both the amount of money and its location in the table of the green room.)

Rakesh further told Bhanwar Lal (25) that he (Bithal Das) used to buy green sarees for Bhanwar Lal's mother when he (Bithal Das) lived in Bombay; and that (26) he (Bithal Das) had been alone when he was electrocuted. All these items were also correct. In addition, Rakesh made one incorrect statement to Bhanwar Lal when he gave (27) his name (presumably of the previous life) as 'Arun'.

According to the informants, Rakesh made all the above 27 statements before the two families could exchange information about the life of Bithal Das and persons and places associated with him. Of these 27 items, four (the name of the previous wife, the mohalla of residence, the name 'Arun', and the activity (electrical wiring at the time of death)) were wrong; all the others were correct.

I should add that informants also credited Rakesh with a number of other statements before and after the two families had met and exchanged information. Most of these were correct, but a few were wrong. In the above enumeration, however, I have concentrated only on the statements that permitted an identification of Bithal Das as the person corresponding to Rakesh's statements.

In view of Rakesh's detailed correct knowledge about Bithal Das, his family and situation, one may reasonably wonder why S.N. Gaur (or the employees of the post office in Tonk whom he first consulted when he reached Tonk) could not find the correct family soon after Rakesh reached Tonk. Instead, they took several hours to do so. The answer probably lies in several facts. First, Bithal Das had died about 20 years before Rakesh came to Tonk, and so young persons would never have heard of him. Second, Rakesh incorrectly 'recognized' (on the edge of Tonk) an electrical pole as being the one by which he had been killed in the previous life. Third, he gave the wrong 'mohalla' (neighborhood) of residence. And fourth, he indicated that he had been

doing electrical wiring at the time of his death.

I think that these three mistakes and the long interval between Bithal Das's death and Rakesh's visit to Tonk put false scents on the trail and that following these consumed some time. I regret that we did not ascertain more precisely just how much time was spent searching for a person corresponding to Rakesh's statements. I do not think, however, that it is correct to say that 'six hours' were spent searching Tonk for the house of the previous life. After his arrival in Tonk considerable time was apparently consumed in useless talk at the post office, were, understandably, no one had ever heard of a carpenter named Bithal Das who had died of electrocution 20 years before. Further time might have been lost if Rakesh (or his father) did not mention the name of Bithal Das in Tonk. Mahi Lal said that Rakesh had not given this name in his presence, but we do not know whether he asked Rakesh to state the name while at the post office. Another employee at the post office at Tonk (the assistant postmaster, H.P. Sharma) stated that during the search for a family matching Rakesh's statements "..... one elderly carpenter said, "certainly there was a person of that name who used to work at the Electricity Board who died of electric shock."" Although the deceased person in question did not work at the Electricity Board, the carpenter's statement (admittedly second hand) is some evidence that Rakesh or his father did mention the name Bithal Das in Tonk. Be that as it may, the knowledge of this elderly carpenter led to the tracing of Bithal Das's family within a short time.

Before I pass to other considerations, I will briefly discuss several of Rakesh's errors. I do this not to explain them away, but to try to understand how they may have occurred. It will be seen that at least some probably arose through processes of association in Rakesh's mind.

a) The name Keshar. Bithal Das's wife was Radha. There was, however, a woman called Keshar living in the compound of Bithal Das's old ('kachcha') house at a time when he lived there himself. We did not learn further details of Bithal Das's relationship with Keshar. I think it worth noting that I would expect a Hindu husband to remember the name of his wife less well than many other details because Hindu men, especially in the area where Bithal Das lived, almost never utter the names of their wives.

b) Chhippa Mohalla. Bithal Das lived in the Kali Paltan mohalla

(ward) of Tonk, not in the Chhippa mohalla. There is, however, a Chhippa mohalla in Tonk, although we did not learn whether Bithal Das had any special relation to it. I learned that chhippas (members of the caste of dyers) did formerly inhabit the mohalla where Bithal Das lived, and possibly some memory of them got into Rakesh's mind. (This detail deserves and will receive further inquiry.)

c) Electrical wiring by Bithal Das. In addition of being a professional carpenter, Bithal Das engaged in a certain amount of electrical wiring, primarily, it seems, as an amateur. He had installed some electrical wiring in his sister's house for a wedding just a few days before his death. Rakesh may have conflated memories of this wiring with the electrocution of Bithal Das that occurred a few days later but had no connection with electrical wiring.

d) The name 'Arun'. Rakesh gave this name to Bhanwar Lal when the latter asked him what his (Rakesh's) name had been 'before'. Bhanwar Lal expected Rakesh to give the name Bithal Das, but Rakesh replied 'Arun'. 'Arun' was a name that Rakesh himself had been given when he had been younger. It seems to me that Rakesh did not understand Bhanwar Lal's meaning when he used the word 'before'.

It remains to add one further comment about Rakesh's errors and the differences in the testimony before and after his first visit to Tonk. I do not agree that the testimony before the visit to Tonk showed disabling inconsistencies and confusions, and that afterward whatever variations had occurred became smoothed out into a correct story given uniformly by all the informants. Rakesh's later statements (after the first visit to Tonk) about the previous life were not purged of errors. For example, he persisted in giving the name of the wife of the previous life as Keshar long after he had met Bithal Das's wife and had had plenty of opportunity to learn normally her correct name, Radha. And although the fact that he did not give the name 'Bithal Das' to Bhanwar Lal, when the latter visited him in Kankroli two days after Rakesh's first visit to Tonk may be regarded as evidence that he did not know the name before he went to Tonk, it could also be interpreted as evidence that he did not learn it when he did go there.

How Rakesh obtained the correct information about Bithal Das

Despite the occurrence of some discrepancies in the testimonies, the informants have convinced me that Rakesh knew a substantial body of

detailed information about Bithal Das before the families concerned exchanged information. If we accept, as I do, the informants' firm statements that they had had no acquaintance whatever before the development of the case, we seem led to suppose that Rakesh had somehow obtained his information about Bithal Das through some paranormal process.

It is more difficult to decide in Rakesh's case than in many others of this type whether he could or could not have obtained this information by extrasensory perception. His father credited him with considerable powers of extrasensory perception. We have not been able to examine the evidence for S.N. Gaur's assertions sufficiently to affirm or exclude the possibility that Rakesh does have the paranormal powers his father attributes to him. Under these circumstances, it seems wiser to close this discussion with the expression of my belief that some paranormal process - not otherwise definable at present - should be considered the most likely source of the correct information that Rakesh showed about the life of Bithal Das.

ACKNOWLEDGMENT

We wish to thank all our informants concerned for their enduring and patient cooperation throughout the investigation of Rakesh's case.

Thanks are due for the support of the Division of Parapsychology, University of Virginia, to: the James S. McDonnell Foundation, the Bernstein Brothers Foundation, and the John E. Fetzer Foundation. We are also thankful to Dr. Ian Stevenson, Dr. J.G. Pratt, and Mr. Champe Ransom for their critical advice. Miss Emily Williams made helpful suggestions for the improvement of the paper.

NOTES

(1) Rakesh's mother, Shanta Devi, was present throughout our two long interviews with S.N. Gaur, but she made no intervention concerning her husband's narration of the events of the case. (This somewhat passive attitude is not uncommon among Indian wives of her situation.) We conjecture that she agreed with what her husband was saying, but unfortunately we failed to question her specifically about this.

(2) In many regions of India, if one asks how many 'bachcha' (children) a person has, most villagers will give the number of sons only. One must ask a second question about the number of 'bachchiyan' (daughters) to learn the total number of children.

(3) H.P. Sharma mistakenly gives the name Keshar that Rakesh gave (incorrectly) for the wife of the previous life. Bithal Das's wife is called Radha.

(4) The phrase 'background entirely wrong' contradicts Bhanwar Lal's statements that immediately follow. He goes on to say that Rakesh had a considerable and accurate knowledge of the background. We therefore think that there was some slip here in the note-taking or perhaps in the translation.

(5) For an example of this type of development, see the case of Gopal Gupta (Stevenson, 1975). The case of Shanti Devi provides another example, although this is not apparent in the original report of her case (Gupta, Sharma, and Mathur, 1936).

REFERENCES

Gupta, L.D., Sharma, N.R., Mathur, T.C. 'An inquiry into the case of Shanti Devi', Delhi, International League, 1936.

Pasricha, S., Stevenson, I. 'Three cases of the reincarnation type in India', Indian Journal of Psychiatry, 1977, 19, 36-42.

Sahay, K.K.N. 'Reincarnation: Verified cases of rebirth after death', Bareilly, N.L. Gupta, 1927.

Stevenson, I. 'Twenty cases suggestive of reincarnation', Charlottesville, University Press of Virginia, 1974. (First published in 1966 as Vol 26 of the Proceedings of the A.S.P.R.)

Stevenson, I. 'Cases of the reincarnation type, Volume I, Ten cases in India', Charlottesville, University Press of Virginia, 1975.

Stevenson, I., Prasad, J., Mehrotra, L.P. et al. 'The investigation of cases of the reincarnation type in India', Contribution to Asian Studies, 1974, 5, 36-49.

Sunderlal, R.B.S. 'Cas apparents de reminiscences des vies anterieures', Revue Metapsychique, 1924, 4, 302-307. Transl. I. Stevenson, 1962.

Division of Parapsychology
Department of Behavioral Medicine and Psychiatry
Box 152, Medical Center
University of Virginia
Charlottesville, Virginia 22908
U.S.A.

PARAPSYCHOLOGY AND REPEATABILITY

Rolf Ejvegaard

My intention in this article is to discuss the claim of repeatability within psychical research. Dr. J.B. Rhine is my example of a parapsychologist who maintains this claim. But, indeed, a vast majority of all scientific researchers in parapsychology do the same. I have chosen Dr. Rhine because he is one of the most well-known in the field and because he has frequently touched upon this topic. I can also from his institute in the United States present an example of unfortunate events that I interpret as consequences of upholding the claim in question.

Before I start my examination I wish to point out that I, in this short paper, do not give all the nuances to the problem. I see it as a pedagogical advantage to formulate the idea of duplication in a rather simple way and to limit the more speculative part of this paper to an absolute minimum.

Almost all scientists think that a serious refutation of parapsychology lies in the fact that so far it has not been possible to show one single experiment that is repeatable at will. For my part, I do not at all think this as being a serious default. My view, as will be shown here, is more or less the opposite one.

The claim of replication has been put forward over and over again by parapsychologists as well as by critics of parapsychology. It has become a 'well established truth' that parapsychology will not be considered a 'real' science until it comes up with the repeatable

experiment.

This 'truth' may be called 'truth by repetition'. When a false statement has been repeated a certain amount of times in scientific journals or text books and has reached the encyclopedias, then this statement attains the nimbus of absolute truth, i.e. a statement that should not be questioned. Such 'truths by repetitions' are very difficult to detect.

Almost every parapsychologist seems to dream of the day when parapsychology in the same way as physics, chemistry, etc. can demonstrate at will its phenomena. These parapsychologists hold that (one of) the most important tasks of their research is to establish the repeatable experiment.

My following point is of no great importance but should be stated to remind us that every science has to be evaluated after its own merits. Would it not be so that on the day we at will can demonstrate telepathy, PK, etc., they cease to be part of parapsychology? Daily replicable telepathy is not parapsychology, it is psychology. The absence of repeatable experiments in parapsychology is then, according to my opinion, not a grave deficiency with parapsychology but one of its presuppositions.

My statement may be looked upon as a negation of an 'established truth'. Therefore, it is, just as this 'established truth', neither deductively nor inductively provable. However, from a pragmatical point of view and to the benefit of psychical research, I consider my statement better than its negation. I hence claim my statement as being pragmatically true.

In the March 1974 issue of the Journal of Parapsychology, J.B. Rhine takes up for discussion, under the headline "Security versus Deception in Parapsychology", some problems of research. His article is 23 pages long. I comment only on a few of his phrases. They are taken out of their context but will, I hope, all the same in a good way mirror the basic attitude with the nestor of parapsychology, namely: The dilemma of parapsychology is the absence of repeatable experiments.

Rhine points out that deception has taken place in all sciences during all times but has become more and more unusual. Also within parapsychology it has become more difficult to cheat. The reason for this, according to Rhine, is threefold.

a) It was mainly the paragnosts or the mediums that in former times cheated. The scientists have acquired better equipment and they are better trained today to check and reveal swindling. This kind of deception, from the paragnost, has almost disappeared.
b) There ia more and more collaboration among the scientists. This has several good consequences among which you find an automatic control system built in, where the researchers are checking each other. An example of this is given later on in this article.
c) The apparatus and devices are getting more and more sophisticated. Specially important is the fully automatic and objective registering of the results. Involuntary mistakes on the experimenter's part can therefore be minimized.

Talking about devices, Rhine adds that "apparatus can sometimes also be used as a screen to conceal the trickery it was intended to prevent.....Machines will not lie, but....." This view later on proved to be of precognitive character.

Rhine also says that "obviously the possibility of easy repetition of tests as a way to check up on a new claim offers the best protection against trickery in research."

Rhine believed (or at least hoped) that tests in parapsychology, if not today then in the future, can be checked by duplicating them. Then trickery would not pay, because it would sooner or later be revealed by other researchers.

According to Rhine, parapsychology has come far towards the reproducible experiment. As an example, he mentions the experiments with clairvoyance that he himself has done and which later have been repeated by others. Rhine adds: "There were also many who, for one reason or another, failed in their attempts to repeat the tests, a fact that means little because there are so many wrong ways to do it." This means that the ones who failed did so because they did not do exactly as Rhine had done. If they had done it correctly then their results would have been positive. This is really saying that the reproducible test exists.

Now, according to Rhine, what kind of researchers cheat? His answer is not very enlightning: The imposters are "weaker minds" or "weaker characters". Many of these "weaker minds" are "able, clever people who need not have used trickery at all; they could surely have learned to do careful, effective testing." In other words, 'clever' persons can

learn to demonstrate paranormal phenomena. Is this really true? Historical data, so far, deny this.

The introduction of the term 'weaker mind' does not explain anything. The weaker mind cheats. How do you know if a mind is weak? Well, a sign is cheating. This is begging the question.

Rhine gives several examples of discharging hoaxing staff from his institution. But, Rhine stresses, this belongs to the past. In the seventies he has been in the position to choose his staff: "Best of all, we have reached a stage at which we can actually look for and to a degree choose the people we want in the field."

Rhine is correct in saying that nowadays there are more persons than ever before who want to do research in parapsychology; but, unfortunately, there is nothing that hinders heads of departments or laboratories from selecting staff who later try trickery. Rhine even adds that "it is impossible for dishonesty to be implemented inside the well-organized psi laboratory today". Surely he is referring to his own lab. This seems more to be wishful thinking than a description of reality.

Rhine takes up two, as he says, important differences between parapsychology and natural sciences. An important difference in this context, I presume, is a difference of kind, not a difference of degree. As an example of a natural science he chooses chemistry. He asserts that it is more difficult to carry out an experiment in parapsychology than in chemistry. Rhine may be right but this is not an important difference. Some experiments both within and between the sciences are more difficult to do than others. This is, however, not a difference in types but a difference in degrees.

The second difference Rhine mentions concerns the amount of researchers. There are many times more chemists than parapsychologists. Rhine is correct again but also here it is only a question of degrees.

From these two differences Rhine concludes the following about the parapsychologist. "If from the start he recognizes independent confirmation to be an essential part of his own goal, he will be able to do much to aid and insure such replication." This is not a logical conclusion from the premisses. It is merely a repetition of what Rhine and most parapsychologists over and over again, but in different

words, come back to: The parapsychologist should strive after the replicable test.

Now, I think, we have to admit that even if it became 'easier' to perform tests in parapsychology and even if the amount of parapsychologists was raised to the same amount as that of chemists, this would, in no way, secure parapsychology to launch the at will demonstrable test.

Rhine continues:".....in parapsychology any new piece of work should be taken as almost a sort of pilot research." Further:".....it is more than ordinary important for another laboratory to repeat it with adequate success. This will then complete the project as a sufficiently effective research contribution."

In chemistry it is possible to carry on tests in an old, drafty house if you do not think that changes in temperature, dampness, smells, draft, house trembling and so forth will effect the experiment. If later on you suspect, for instance, that the moisture in the air had some effect on the result of the experiment, then you need only to do it again with this factor under control. You may, if you please to do so, call the first experiment a pilot study and the second one the final experiment.

Rhine reasons in this same manner as regards to parapsychology. I think he is wrong in doing so. A pilot study in parapsychology is for trying out new methods, equipment and material, to see that you handle the test situation or to get acquainted with it, not to evaluate its fruitfulness from the test results. This means that when you in parapsychology go to the 'real' experiment it has to be planned in every detail. Nothing can depend on chance. All possible conditions have to be under control. The experiment can neither be verified nor falsified by other tests. Rhine's and my views on this point are thus imcompatible. I should add here that a test with all conditions under control automatically fulfills the claim of duplication provided the nature of the variable investigated is replicable.

Dishonesty cannot be explained by 'weaker minds'. Instead the experimental atmosphere plays an important role. One way of diminishing cheating, therefore, would be to change the research atmosphere for the parapsychologists. When they are not obliged to hand in positive results or come up with the repeatable experiment in order to have their findings published (and here we touch upon

publishing policy) or their contracts prolonged, then they are not lured into diddling any longer.

Shortly after Rhine wrote that it was impossible (sic) for dishonesty to be implemented inside a well-organized lab of his own type one of his fulltime researchers was disclosed as an imposter.

Levy was a young, very successful researcher in parapsychology. From 1971 to 1974 he published about a dozen articles and became during that time the most productive in one of the most well-known and respected journals of parapsychology in the world. He seemed to be on his way to find the replicable experiment within anpsi.

While Levy was carrying on with his animal precognition and PK experiments one of his assistants noticed that Levy unnecessarily often checked the registration device, which automatically filed the results. Three of the assistants of the lab decided to check what Levy was doing and parallel to Levy's device, without him noticing it, they introduced their own registration apparatus. The assistants only reached results in accordance with chance expectancy on their equipment while Levy got positive, highly significant results on his.

Levy, without any fuss, admitted to Rhine that he had been cheating and he was forced to resign. Hence, Levy became a new example of a researcher that had been fired by Rhine.

Why did Levy, a highly respected man who got his medical doctor's degree from the Medical College of Georgia in 1973, get involved in this unwise manipulation? Unwise, because Rhine's second reason (see above) for the diminishing of deception in sciences, the built in control system with researchers checking each other, really worked beautifully in this case. He himself has given an answer and I am ready to accept it. After close to fifteen successful experiments the positive results dropped. The longer it took for positive results to reappear the more he felt the pressure to do something radical. Everyone, including himself, expected positive results. He was the one who seemed to, finally, be near the repeatable test.

After the unmasking of Levy Rhine declared that all the experiments by Levy must be considered invalid until they had be verified by later experiments. Rhine's line of reasoning would have been correct if it had been a question of chemistry. But when it comes to parapsychology every test has to be deemed by its own merits. Of course, you may

still come to the same verdict as Rhine.

For my part, I think Rhine's conclusion that Levy cheated in all his experimenting was premature. That Levy used trickery in his latest tests does not necessarily mean that he did so in his earlier ones. Paragnosts have sometimes admitted that they, in order to get the participants in a favorable mood, use trickery at the beginning of an experimental situation. From the fact that a paragnost, in cases where he himself decides the test circumstances, sometimes is hoaxing, it does not follow that he always does so.

In the Levy case, it is even, considering psychological circumstances, more probable that the earlier results were genuine. If a medical student (as Levy during the sixties) with an interest for parapsychology gets negative results in his experimenting, then he hardly would continue to put in all his spare time in dishonest experimenting. Had it shown that he belonged to the group of experimenters (and there seems to be such ones) that hardly ever gets positive results, then he most likely would have shrugged his shoulders and put in more time on his medical studies. He would certainly not have tried to make experimental parapsychology his career.

On the other hand, after becoming the chief researcher at the most famous parapsychological lab in the world and the man who most likely would succeed Rhine there was quite a pressure on him to produce positive results.

The demand that a researcher continuously should show new and good results is damaging to parapsychology. The consequence, among others, is a higher degree of deception than otherwise. Upholding the claim of test duplication will lead to new Levy cases. If you instead reckon that paranormal ability is cyclic, coming and going, then the researcher expects positive results to stop soon. To put all the burden for what happened in the Levy case on Levy alone almost looks like a way of freeing other culprits.

Instead of applauding when a researcher is getting an unusually long series of successful trials one should begin to be suspicious. Then there may be something else than a paranormal phenomena involved. In principle this means that you have to be on guard already at the first positive result and that the suspicion is of an additive type, i.e. that it becomes stronger the longer the sequence with positive results

gets. The suspicion does not mean that the experimenter is fraudulent, but that something else, whatever it may be, not paranormal, effects the experiment.

The sooner the parapsychologist in his thinking can get away from believing that the repeatable test is benefiting and starts to realize its disadvantages, the better off parapsychology as a science will be. It then might be easier to work out a suitable frame-work for psychical research. As it looks now, parapsychologists are blocked. They will never master the repeatable experiment and do not understand how to go on in the field without it.

Is it not the opponents of parapsychology who are always reprimanding parapsychologists for not being able to come up with the at will demonstrative paranormal phenomenon, that have lured the parapsychologist into the swamp where they now are sitting fast? Better would be if parapsychologists did not listen to persons unfamiliar with the field and instead began to develope the science after its own prerogatives.

Now, before ending this article I want to, very briefly, touch upon two problem complexes.

I. What is science and does hypothesis testing imply repeatability?
II. Is it even thinkable that there are sporadic events and, in case of 'yes', how come?

There is, to start with question I, no consensus among scientists as to what the meaning of science is and as, for instance, Feyerabend ('Against Method') has pointed out there are no generally accepted principles to decide if an experiment is to be considered scientific or not. There are criteria for distinguishing non-scientific activities. None of those are connected with repeatability.

Researchers in established sciences try to uphold the claim of hypothesis testing or testability. Of course, it is not easy to do so in social sciences as psychology, antropology or political sciences. There is no simple experiment to be carried out to test the hypothesis of the Neanderthal-man and the Peking-man having the same ancestors.

My question is: Does the claim of testability imply experiment duplication at will? To answer that question we have to realize that we can distinguish among four different kinds of testability (see figure 1).

FIGURE 1

Testability types	Verification	Falsification
Practical	Type A	Type B
Theoretical	Type C	Type D

Pro primo, we have the orthodox claim that a hypothesis must be testable in reality and the modern claim that it is enough for the scientist that he in theory can give instructions for the testing of a hypothesis. In this latter case it virtually means that there is no actual possibility to put the test into practice.

Pro secundo, the testing may be carried out as a verification or as a falsification of hypotheses. The latter kind can be regarded as the stronger (or more challenging) type of testing a hypothesis.

Testability of types A and B imply repeatability. In other words, phenomena which are not demonstrable do not fulfil testability of types A and B. One could even say that reproducibility at will and testability of type A are one and the same.

Testability of types C and D does not imply duplicability. According to these types it is sufficient to outline a method which in principle could be used for testing the hypothesis. Thus, hypothesis testing in principle is of no objection to psychical research.

To clarify matters, I will give an example of a testability case of type C. In a test situation Lapp children aged seven from the village

V1 learned the multiplication tables 50% faster than the seven year old Lapps in V2, mutatis mutandis. A hypothesis stating that this difference depends on the death of the mathematics teacher a few days preceeding the test can not be tested according to testability A or B. Testability C and D remain for us. Even if it is far fetched, we could imagine that the test situation was built up anew. I leave it to the reader to do this reconstruction in theory.

I should add here, and my example with the Lapps illustrates it, that it sometimes is debatable if a certain experiment fulfills the claim of testability C or D. This means that the notion 'testability' is, as so many scientific terms, somewhat vague.

Testability in theory, and here we pass over to my question II, discloses the real sore part of parapsychology: the lack of a general theory or even frame-work.

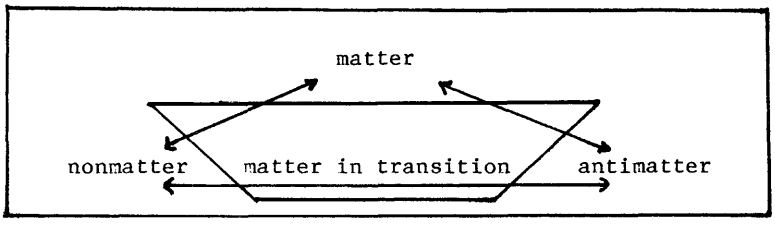
If one suggests that there are phenomena, which are not repeatable at will, sporadic phenomena which are less stable, less categorical than normal phenomena, i.e. the ones we meet in daily life and study in physics and chemistry then one should also be able to point out the direction where to find the theory for such a claim. I conclude my paper by giving an outline of such a theory.

Let us assume that there is matter, antimatter and nonmatter. Antimatter is just a term that I have borrowed from modern physics and which I leave undefined here. With nonmatter I simply mean emptiness, nothingness. Let us also assume that, in contrary to the ideas of classical physics, matter is not constant, not eternal, but that bits of it can be created out of nothing and be destroyed to vanish into nothing.

Now, in the world of stable matter all events will be stable and within certain limits repeatable. Unstable events occur in the world outside the stable matter, in the worlds where matter is in transition from one kind to another (see figure 2).

From the assumption that matter can change from one form into another form there follows that some kind of connection between matter and nonmatter, between matter and antimatter, and between antimatter and nonmatter exists. Hence, there must be some kind of connection between events in matter and in matter-in-transition. The sporadic events then could in very rare cases influence the normal events and

FIGURE 2



Processes:

from nonmatter to matter
from matter to nonmatter
from nonmatter to antimatter
from antimatter to nonmatter
from matter to antimatter
from antimatter to matter

Phenomena:

in the matter area: normal
in the antimatter area: normal?
in the matter-in-transition area: paranormal
in the nonmatter area: none

make the latter look less normal. This could be the mechanics behind the observed paranormality in our world.

Maybe it is not necessary, but for all eventualities, I will admit that I do not think my model is the one that will be accepted by science sometime in the future. The purpose of my sketching it is only to show that there are ways of including paranormal events of all types in a scientific paradigm.

SUMMARY AND CONCLUSION

Parapsychology is a science in which one applies means and methods used in natural and humanistic sciences. It can however never be a question of having a duplication method in parapsychology as one has it in let us say chemistry and physics. If parapsychological experiments were reproducible at will then there would be no science called parapsychology. The very day all parapsychological experiments were repeatable with highly significant results, that day the different parts of psychical research would disappear into physics, psychology, sociology, and so on.

If the repeatable experiment does not exist in parapsychology this in no way can be regarded as a disadvantage or a lack or a misfortunate, but one of the very fundamentals for parapsychology and something we have to live with. Once this simple fact is grasped it inevitably has to influence the planning, the screening, the setting up and the carrying out of parapsychological experiments. Pilot studies then exist only in so far as trying out new techniques are concerned. When it comes to the results and evaluation of parapsychological experiments there are no pilot studies what so ever. This means that carelesness and lack of planning and thought never can be rectified or corrected afterwards. This indeed makes experimenting in parapsychology a most difficult task.

Curing parapsychology from the duplication stigma would help to clean up the fraudulent tendencies in psychical research. No one would any longer feel compelled to come up with the repeatable test.

Provided there are paranormal events we have to draw some conclusions about the world and world order. Certainly the world is not 'that' orderly as scientists of all kinds choose to think. The world is maybe not at all 'that' fixed, stable or regular as we hope for or as we make it seem to be when we, mainly for pedagogical reasons, set up rules and formulate natural laws. This way of handling facts instead of helping us many times perhaps hinders us and makes us blind for what there really is. Not knowing exactly what we are looking for in parapsychology and applying methods from physics, maybe has taken us into stray ways. Perhaps today we are not setting up experiments that are revealing but concealing paranormal events.

Psychical research points towards the direction that there are

events of different degrees of reality: Those we learn about in school and that can be demonstrated at will; those that parapsychology has started to trace and which are of a 'lower' degree of existence. It is possible (see figure 2) to draw a model in which these two kinds of events interact.

REFERENCES

Ejvegaard, R. 'Parapsykologi som vetenskap', Sokaren, 1974, No. 9.

Ejvegaard, R. 'Replik', Sokaren, 1975, No. 2.

Ejvegaard, R. 'La repitibilita in parapsicologia', ESP. Mensile di parapsicologia e fenomeni dell insolito, 1976, No. 16.

Ejvegaard, R. 'Vetenskapsteori och parapsykologi', Varld och Vetande, 1978, No. 3.

Feyerabend, P. 'Against method', Minnesota studies in the philosophy of science, 1970, Vol. 4.

Johnson, M. 'On publication policy regarding non-significant results', E.J.P., 1975, Vol. 1, No. 2.

Murphy, G. 'The problem of repeatability in psychical research', J.A.S.P.R., 1971, Vol. 65.

Nilsson, I. 'The paradigm of the Rhinean school. Part I and II', E.J.P., 1975 - 1976, Vol. 1, No. 1 and 2.

Rhine, J.B. 'Security versus deception in parapsychology', J.o.P., 1974, Vol. 38.

Rhine, J.B. 'A new case of experimenter unreliability', J.o.P., 1974, Vol. 38.

Rhine, J.B. 'Second report on a case of experimenter fraud', J.o.P., 1975, Vol. 39.

R. Ejvegaard
Mellansjovagen 35
S-14146 Huddinge
Sweden

THE REPEATABILITY OF SIGNIFICANCE AND THE SIGNIFICANCE OF REPEATABILITY

Carl L. Sargent
University of Cambridge

The terms 'significant' and 'repeatability' (or 'replicability') are often used in parapsychology without a precise specification of what they mean. I shall suggest that both are polymorphous concepts and that distinctions between different types of significance and repeatability are exceedingly important. More precise use of terms would help parapsychology.

SIGNIFICANCE: THREE TYPES

First there is statistical significance. I shall assume that the reader of this Journal knows exactly what this means although a good reference on this point is Siegel (1956). I shall point out one aspect of statistical significance which may seem trivially important but which may influence assessments of repeatability (which I shall discuss later).

Conventional criteria of statistical significance - whatever they are, the .05 level, the .01 level, or even the .10 level sometimes used in psychology - must be recognized as completely arbitrary. It may seem that I am stressing the obvious when I point out that P-values of .40 and .051 are very different, even if they are both 'insignificant' if alpha was set at .05, and that .049 and .0001 are

also very different, even if they are both significant when alpha is set at .05. The 'significant'/'insignificant' dichotomy is an arbitrary one superimposed on a natural continuum. If one attempted to replicate a previously observed statistically significant finding and obtained a P-value of .051 for the result, is this really insignificant and a failure to replicate? Surely we would not fall over ourselves to proclaim a failure to replicate if we obtained such a P-value in such an experiment, especially (to anticipate the next type of significance) if the magnitude of effect was similar to that observed in the first study we were trying to replicate.

Now this issue affects repeatability and assessments thereof. Consider the data in table 1, which are mythical. If we assess

TABLE 1
Outcomes of eight similar experiments; Imaginary data

Experiment number	P-value outcome	Chi-square (df=2 each case)
1	.0018	12.64
2	.007	9.92
3	.065	5.47
4	.07	5.32
5	.095	4.71
6	.11	4.41
7	.18	3.43
8	.40	1.83

Sum chi-square = 47.73, df=16, P<.001

repeatability in terms of the arbitrary dichotomy, significant or insignificant, setting alpha at .05, then only two of the eight experiments in the table - we shall assume that they are very similar and their data can be pooled to assess repeatability - show a significant result and this is not itself significant. The exact probability of obtaining two or more 'hits' in eight 'trials' with

p(hit)=.05 is .057. However a superior statistical procedure is to utilize a statistic which dumps this unnatural dichotomy, like Guilford's (1956) chi-square test where

$$\chi^2 = \Sigma -2\log_e P \quad \text{with } 2N \text{ df}$$

Application of this statistical test to the data gives Chi-square is 47.73 with df=16, P<.001, which tells us that there is a consistent effect in the data. Use of this test (a) abandons the unnatural and arbitrary dichotomization of results around an arbitrary alpha level and (b) is more sensitive than a dichotomy-based evaluation of data.

Now whether one should use the Guilford Chi-square or analogous statistics, or the more common dichotomy-based analyses of repeatability, is neither here or there for the moment. The point I want to make is that criteria of statistical significance are arbitrary; we should not forget this; and that too strict an adherence to conventions relating to statistical significance may be short-sighted and insensitive.

Finally strict adherence to conventions relating to statistical significance may waste time. In pilot experiments of a novel kind one may be wasting time running a moderately-sized experiment which one feels has a large enough sample of subjects/trials to get a significant result out if the hypothesis one is working on is 'correct'. It may be wiser to use a small-scale study approach to seek insignificant trend confirmation of one's ideas. This saves time.

Now to a second type of 'significance', what one might term practical significance or the magnitude of the effect. Clearly large-size psi effects are preferable to small-size effects. This is increasingly being recognized (e.g., Honorton, 1977; Tart, 1978). For one thing, small-size effects are much more likely to be artifacts than large-size effects (but see Sargent, 1980a, Sondow, 1980). For another, scientists outside of parapsychology are very much concerned with this factor (e.g., Stevens, cited in Ransom, 1971; Evans, 1979). In my experience statistical significance is a precondition for interest. If small P-values can be demonstrated for experimental outcomes in well-controlled psi experiments the next question asked by broadly sceptical outsiders has been, in my experience, "what is the magnitude of the effect?". The great interest in the exploits of Geller on the one hand, and the remote viewing work of Targ and Puthoff (1979) on the other surely has a great deal to do with the

fact that the psi effects reported are of large magnitude.

Once again it may seem that I am stressing the obvious, but this is not so. In a recent paper, Pratt (1978) stressed the need for more 'good evidence' in parapsychology. What this 'good evidence' was in Pratt's view was not clearly stated in that paper but Gaither's other writings make it clear that he was thinking in terms of well-controlled studies, probably with selected subjects, which showed moderately consistent and very highly improbable extra-chance effects. Nowhere in this formulation did Pratt ever mention magnitude of effect. I did not have the chance to get a reply from him on this matter before his untimely death; but even if he would have agreed with me, the fact that he did not mention the magnitude factor in his papers testifies to a lack of awareness of the importance of this factor.

So large-magnitude psi effects (a) are less likely to be artifacts than small-magnitude effects and (b) are more likely to impress outsiders. Finally they are more likely to be replicable, to be robust in the face of the inevitable error variances associated with inter-laboratory and inter-experimenter differences.

The third type of significance I refer to is theoretical significance. Stanford (1974c) has provocatively commented that a theoretically interesting finding significant at the .05 level is more important than a theoretically boring one significant at the .01 level. I believe that this is correct, and in keeping with the whole historical development of parapsychology for the last 40 years. Eysenck stated many years ago (1957) that unless there was a large-scale conspiracy afoot psi was a genuine phenomenon and, adding the evidence collected in the 23 years since, one can hardly deny this. Our attention has been turned increasingly to trying to understand psi, to detail and theorize about those conditions, psychological and physical, which influence psi events. Mere 'demonstration experiments' tell us precisely nothing about psi, and no more do recherche effects in psi experiments which lie outside of explication, even if they are highly significant and replicable. This point was hammered home to me when, in 1978, I conducted a large-scale test using the Names Test (Sargent, 1977) for psi and observed a significant overall psi-missing effect (P=.02). Since none of the subconditions gave differential effects with respect to each other I had no idea why this happened. Although the overall data gave significant evidence of psi - and hence two of the four experiments on

this theme (Sargent, 1977, plus the 1978 study) had given significant missing at the .02 level, I did not know why. The significant-statistically-psi-missing told me nothing. The finding was statistically significant but theoretically insignificant.

Later I shall comment on the importance of theoretical frameworks for repeatability in some detail. At this point let me state that theoretical significance of results is seen as vitally important by scientists outside parapsychology. Most excitement is, naturally, caused by a theoretical structure such as the Walker (1975) / Schmidt (1976) models which suggest a possible natural explanation for psi, but Stanford's elegant PMIR model (Stanford, 1974a, 1974b, 1977b), illustrated with the PMIR pursuit rotor tracking study (Stanford, Zenhausern, Taylor, and Dwyer, 1975) is also received well.

REPEATABILITY

Consider the following sequence of events. An experimenter tests the hypothesis that performance in a psi task is correlated positively with variable Z. He tests this hypothesis, evaluating it by correlating Z and psi measures with a Pearson (or Spearman) statistical measure of correlation, and finds a positive correlation with a value large enough to reject the null hypothesis in favour of his experimental hypothesis at the preset alpha level (usually .05). He now attempts to reproduce the conditions of the original experiment, runs another bunch of subjects, and once again rejects the null hypothesis. We say here that the two experiments have both given 'significant results' and that the second experiment replicates the findings of the first.

Such a state of affairs we may refer to as intra-experimenter concrete replication; intra-experimenter because the same experimenter ran both experiments and concrete because the attempt to reproduce the original experiment in precise detail was made. This is not, in fact, what people generally mean when they refer to 'repeatability' in the abstract; they generally refer to the inter-experimenter case, i.e. when two different experimenters run the two experiments.

It is commonly asserted that inter-experimenter repeatability is much lower than intra-experimenter repeatability (e.g., Beloff, 1973). It is useful to point out that this will be expected to some extent

for the two following reasons:

(1) Clearly, variability between studies across time conducted in one laboratory is less than variability between studies across laboratories. Factors which enter the picture here would include (a) subject population differences (a dismally unexplored factor), (b) set and setting factors (Terry and Honorton, 1976), and so on.

(2) An experimenter running experiments on a new theme is most likely to publish his results if both a pilot and confirmatory study(ies) are significant. If this outcome does not result he may say, very reasonably when new work is undertaken, "Well, my novel hypothesis may well have been wrong, It is premature to present the data." Because of this stance - and it must be stressed that this is quite reasonable when a novel hypothesis is being explored - the repeatability for this intra-experimenter case will appear to be very high over numerous experiments of this type. Now, subsequent studies by other experimenters are much more likely to be offered for publication (and accepted) with null data than is the case with the original intra-experimenter case, because these subsequent experimenters are publishing replicative data, i.e. data from experiments in which they were attempting to replicate the work of someone else. They do not have the grounds for nonpublication of null data which the original experimenter had. Hence, artifactually, inter-experimenter replicability will appear to be lower than intra-replicability. The artifactual difference arises from the fact that the original experimenter will be less likely to publish null data than subsequent experimenters. Indeed it is not unreasonable for him to do this as we have seen. It is not, of course, reasonable for him to not publish subsequent failures to replicate on his part. Experimenters do not tend to do this; indeed the literature on within-experimenter declines testifies to this (Kennedy and Taddonio, 1976). I am personally inclined to think that many of these within-experimenter declines are artifacts (Sargent, 1980b).

To correct the record, however, there are of course many instances one might give of experimenters not being able to replicate their own results whilst other experimenters have been able to replicate them. Examples would include Nicol and Humphrey's (1955) inability to replicate their extraversion/ESP data (Nicol and Humphrey, 1953), and subsequent successes with this finding (Sargent, 1981); failures with some dream telepathy studies at Maimonides (Honorton and Ullman, 1975) which have been replicated elsewhere (e.g., Child, Kanthamani, and Sweeney, 1977); Stanford's difficulty in replicating some of his response-bias effects (e.g., Stanford and Schroeter, 1978) in contrast

to other successful work of this type (e.g., Sargent, 1977b, 1979, 1980c); occasional 'hiccups' in the Maimonides ganzfeld work (Honorton, 1978a) and the successful replication work elsewhere (e.g., Sargent, 1980b; Schmitt and Stanford, 1978), and so on.

Now let me turn to a different dichotomy within repeatability, the distinction between concrete and conceptual replication. I am flattered by the attributions of this distinction to me made by Braud (1979) and Gruber (1979) but the distinction was, of course, originally made by Stanford (1974c) most explicitly. A concrete replication scenario we have already discussed. The conceptual replication would be exemplified by the following sequence of events. First an experimenter tests hypothesis A, which states that psi should be positively correlated with variable Z, by manipulating conditions so as to optimize Z and hence optimize psi if his hypothesis is correct (for an example, for Z read relaxation and for experimenter read Braud and Braud, 1973). This works out. Now the experimenter takes a new line. Rather than replicate the experiment in concrete terms, manipulate the conditions so as to depress Z and see if you get significant missing. This, of course, is just what Braud did (Braud and Braud, 1974). The second experiment clearly confirms the hypothesis tested by the first experiment even though the concrete details of the two experiments are not the same.

Now I believe that concrete replications are greatly inferior to conceptual replications for one major reason; concrete replications do not favour progress in hypothesis-testing. No hypothesis predicts only one experimental outcome and thus conceptual replications - testing different predictions about outcomes of experiments - must be preferable to mere concrete replications. Further, concrete replications are most likely to fall foul of nonspecific effects than conceptual replications. Tart (1977) has suggested that control groups in experimental/control group comparisons may score poorly in psi testing simply because they are control groups; one could never detect such systematic error variance in concrete replications but one can get to grips with this problem in the context of conceptual replications.

The other good reason for not overdoing it on concrete replications is that a sequence of such experiments is profoundly boring. Experimenter boredom can communicate itself to subjects through sensory channels and inhibit psi-elicitation. The Duke group realized this a long time ago. Experimenter boredom would also hardly be

expected to favour experimenter psi effects, if they exist.

If it appears from my final comments here that I am inclined to agree with the suggestion of White (1979), who has suggested that experimenter motivation is the key to obtaining significant evidence of psi, far from it; there is not a shred of respectable experimental evidence in favour of such a view. It just does seem moderately plausible that boredom will not help psi-elicitation, and I don't enjoy being bored; nor, I should have thought, do many parapsychologists.

ASSESSING REPEATABILITY

This is a key issue in parapsychology. Many important arguments are involved here. Let us start with Honorton's 'box score' analysis of repeatability of, e.g., ganzfeld psi-optimization (Honorton, 1978a) and psi-optimization in microdynamic PK studies (Honorton, 1978b).

Honorton collates all published studies and evaluates the number which reject the null-hypothesis at the .05 level and/or computes a chi-square of the Guilford type for all studies. When exact P-values for outcomes are not given by authors of papers Honorton assumes P=.99, which is rather conservative; I should have thought a good case could be made out for P=.50, but certainly Honorton's approach is conservative on this count. Now such a procedure has obvious advantages. For one thing hot air is replaced with a data summary. Further these data summaries (for example, Honorton's survey of published work between 1934-1939 inside and outside Duke; Honorton, 1976) sometimes provide surprising results.

Now let us examine the procedure itself and see what comments might be made about it.

(1) Experiments or experimenters?

Millar (1979) criticizes the microdynamic PK survey (Honorton, 1978b) on the grounds that analysis by experimenters is a more valid approach than analysis by experiments. He presents a re-analysis of Honorton's data classified by 'psychic' and 'non-psychic' experimenters (using a 2x2 chi-square with nonindependence of observations, incidentally) to support his thesis relating to

experimenter psi effects. A fair comment would be that both survey methods are valuable. The analysis by experimenters case is useful as a secondary analysis to the type favoured by Honorton to check if different repeatability figures are given by the two analyses. If the figures are startingly different one's conclusions would be appropriately influenced.

(2) Selection of data - Yes or no?

Honorton, in his 1978 survey (but not in his 1976 survey) did not eliminate any data on any grounds from his survey. Palmer (1971, 1972, 1977) has eliminated some data from surveys on the sheep-goat, extraversion, and neuroticism correlates of psi on statistical grounds. There is a major problem here.

If one takes all the data from experimentation on, say, ganzfeld psi-optimization, then one is indeed, as Kennedy (1979a, 1979b) says, mixing wheat and chaff. Some ganzfeld studies are not truly 'replications' at all; e.g. Rogo's (1976, study I) study in which no auditory ganzfeld was used and his studies (1976, II and III) in which the duration of ganzfeld was certainly insufficient to optimize psi if Honorton's (1977) theoretical scheme is correct (null results were indeed obtained). Several other studies appear to have had sensory cueing possibilities (see Sargent, 1980a; Sondow, 1980; Palmer, Khamashta, and Israelson, 1979). Probably the answer is to analyse data (a) altogether and then (b) selected with respect to specific criteria such as the possibility of sensory cueing (Honorton, 1979) and the duration of ganzfeld exposure in that concrete example (Honorton, 1977).

(3) Dichotomies or other analyses?

One may end up from a data survey by saying, i.e., 8 of the 16 published studies of this type gave significant (at the 5% level) outcomes, the repeatability level is 50%, ten times the chance error rate, $P<.000001$; or one may prefer the chi-square type analysis referred to earlier which is certainly more sensitive. As Palmer (1977) has pointed out, with a genuine effect, trend data from statistically insignificant findings should corroborate imbalances in statistical results between one pole and another (e.g., negative and positive correlations between psi and variable Z) and thus the chi-square is a more useful measure.

(4) Selection of data revisited

Now to a more problematic issue. It is often claimed, even by some parapsychologists (e.g., Parker, 1978) that much insignificant data is unpublished and hence surveys of published data may give wildly misleading impressions of repeatability. Two important points can be made here at once. The first is that the incidence of publication of null results has undeniably increased in the last 25 years and yet the repeatability level of the best recent lines of research, like the psi-conducive states, is certainly higher than that reported for the old semi-consistent predictor variables like attitude and extraversion. This suggests very strongly that a real improvement in repeatability has been seen in parapsychology during this time. A second, related, point, made by Honorton (1978) with respect to the ganzfeld work, is that there are not very many parapsychologists around, the studies are time-consuming, and that since 1973 nearly 40 studies have been published of which around one-half have given significant results. It seems inconceivable that in such a short time the vast number of chance data-sets needed to cancel out this apparently highly impressive track record could have been done.

It is not however necessary to rest our case here. Another argument against selective publication policies rendering empirical assessments of repeatability largely worthless (which is in itself an unscientific argument since it appeals to a nonmeasurable variable) will be given a subsequent section to itself since it completely eliminates the objection. Finally it may be pointed out that the reply to the smug sceptic who says that selective reporting makes appearances of repeatability illusory is to reply "Maybe, but for all we know all those unpublished experiments may have had hopeless methodological failings. Indeed, since they have not been through a process of peer review in journal refereeing (or have been and were rejected), this seems a not implausible possibility". Impasse.

Very well. We have now to consider what we are going to say when we get the results of our box-score analysis. Let us assume we find a 50% repeatability rate for an experimental finding. Is this good, is it bad? What criterion are we to measure this by?

One criterion I think we can all reject is that an experimental finding should be perfectly repeatable, i.e. an experiment can be repeated over and over again and the null hypothesis will be rejected in favour of the pre-stated experimental hypothesis at the preset

alpha every time. This just doesn't happen in social sciences. I suspect that it doesn't in physical sciences either, but that's another issue; at any rate, since I have never heard anyone seriously defend the perfect repeatability criterion, I shall not discuss it further.

Honorton's (1976) suggestion is that we should take empirical repeatability in psychology as a yardstick. If findings in parapsychology are broadly as repeatable as those in psychology, then parapsychology is doing well, or as least as well as can reasonably be expected.

Although this seems hard to argue against, some people have in fact done so, most recently Broughton (1978). Broughton argues that we need higher repeatability in parapsychology than exists in psychology because of the a priori less likely nature of the phenomena studied in parapsychology. Now for one thing a priori arguments are pretty dubious things; but in any case the point can easily be refuted. It is true that high repeatability is better than low repeatability so far as 'evidence for' is concerned but it is not necessary for parapsychology to have higher repeatability than exists in psychology for stronger 'evidence for' to be obtained. A larger number of successes with repeatability rate identical will do. The exact probability of two or more experimental results significantly (at the 5% level) supporting an experimental hypothesis out of three attempts is .0071. The exact probability of four successes out of six of this type is <.00001. The latter case is clearly stronger 'evidence for' than the first but the repeatability rate is equal in the two cases. So Broughton's point is clearly erroneous. Parapsychology does not need higher repeatability than psychology and indeed it is grossly unreasonable to expect human beings to be more predictable in psi experiments than they are in psychological experiments.

Now the next step is to decide what area of psychology we might best compare our repeatability with. It seems to me that subliminal perception and areas of research into weak sensory systems, e.g., the olfactory system, provide reasonable parallels. What we find if we look there is (a) that repeatability is hardly assessed at all, as is the case throughout psychology (Barber, 1976) and (b) broadly, the data are a mess. Taking the olfactory system for example, in a recent authoritative review by Moskovitz (1978) I could not locate the word 'repeatability' anywhere and the old cliches of "more research is needed" and "other variables may interact here" (surely readers can

supply other cliches for themselves) were in abundance. Indeed in some of my own work with the ganzfeld technique (Sargent, 1980b) I have had more trouble in replicating psychological findings than with replicating parapsychological ones.

THE DIRECTIONAL ANALYSIS OF SIGNIFICANCE PROCEDURE (DASP)

Before concluding this paper I wish to eliminate completely arguments relating to selection of data. We noted previously that surveys of repeatability can be attacked on the grounds that nonsignificant outcomes are not always published (neither are significant ones, of course; I have six unpublished significant response-bias effects in my filing cabinets) and therefore the surveys are misleading. We have also seen that this is an unscientific and disreputable argument. However we can circumvent it completely. Whilst we cannot assess repeatability in the traditional sense with the DASP we can show unequivocally that a real effect exists in data sets.

The logic of the DASP comes from Palmer (1977). Palmer noted that if no real effect of a variable exists on psi, then taking significant outcomes alone one should find an approximate balance between positive and negative correlational findings. If no real relationship exists between ESP test performance and extraversion, for example, if we stick to the reported significant effects alone there should be an approximately equal number of positive and negative sign correlations. The beauty of this statistical analysis is that it is completely unaffected by suppression of null data.

Now let us use the DASP on certain experimental findings in parapsychology and see what happens (see table 2).

We may see that the existence of real effects of the types noted is quite clear in the data. The DASP shows clearly that psi-hitting is elicited by ganzfeld stimulation, the dream state, and progressive relaxation; that extraverts and low-neurotic persons score higher on ESP tasks than introverts and neurotics; that low-bias responses are higher scoring than high-bias responses in ESP tasks and so on. There is some 'rubbish', of course, but this is to be expected (it would, in fact, be extremely worrying if there was no 'rubbish' at all).

The DASP shows effects in data sets and is invulnerable to the

TABLE 2
The directional analysis of significance procedure applied to some hypotheses in parapsychology

Hypothesis	Significant confirming instances	Significant negating instances	Exact binomial P two-tailed
1. Believers in ESP score higher in ESP tests than disbelievers (Palmer, 1971, 1972, 1977)	17	1	.00014
2. Extraverts score higher than introverts in ESP tasks (Sargent, 1981)	16	0	.000031
3. Low-bias responses show higher scoring in ESP tasks than high-bias responses (Stanford, 1967; Carpenter, 1977; Sargent, 1979)	16	1	.00026
4. Ganzfeld isolation elicits psi-hitting rather than psi-missing (Honorton, 1978; Sargent, 1980b; Child and Levi, 1979; Keane and Wells, 1979; Sargent and Matthews, 1980; Lane, Matthews, Radcliffe, Sargent and Harley, 1980)	19	2	.00018
5. The dream state elicits psi-hitting, not psi-missing (Ullman, Krippner and Vaughan, 1973 and subsequent research)	10	0	.002
6. Scores on the Defense Mechanism Test correlate positively with ESP test performance (Haraldsson and Johnson, 1979)	8	0	.0078
7. Low-neurotic subjects score higher than high-neurotic subjects in ESP tasks (Palmer, 1977; Sargent, unpublished data)	9	0	.0039
8. Progressive relaxation exercises elicit psi-hitting, not psi-missing (Honorton, 1977; subsequent studies)	12	0	.00049

Note: Data base used in most cases is unselected, total data from English language publications. In some cases the original compilers of the survey did select data on some grounds (e.g., Palmer, 1977, deleted some data on statistical grounds, although this led to a more conservative estimation of the significance of results than would otherwise have been the case).

effects of selective reporting unless one argues that experimenters actually suppress findings significantly opposite to the prevailing concepts which, in parapsychology of all sciences, is a bizarre argument (as well as being one capable of demolishing every scientific finding in existence). The DASP is in one sense a highly sensitive measure and suggests that with directional hypotheses we can forget all about null data completely in our epistemology.

FINAL COMMENTS

I have not attempted in this paper to raise cogent arguments about repeatability and significance. My own views are not sufficiently clear for that, and I prefer to raise what I feel to be interesting, under-debated, and provocative points to stimulate thought in these difficult areas. There is, however, one thing about which I feel very strongly. Somewhere in the recesses of my memory (which may fail me here but that isn't too important) I can recall Larry LeShan citing Henry Margenau to the effect that "I wish you guys would sometimes stick up for your data". I feel that we are far too ready to extol our shortcomings in parapsychology, that we are far too critical in some ways. For some failings, such as permitting sensory cueing possibilities of a fairly obvious kind or measuring independent variables after subjects have had feedback on ESP performance and proceeding to correlate independent and dependent variables, no criticism is too strong. We really should have learned not to do stupid things like that by now (or at least not to publish them!). However, many times in parapsychology a promising line of experimentation has been prematurely killed because of a couple of 'failures to replicate'. Let me contrast that with two sequences of events in 'normal science'.

Collins (1974, 1978) has given observations on the attempted replication of the TEA laser, a device first constructed in Canada in 1968 and of which details were made public in 1970. The first successful 'replication', i.e. a functional TEA laser built elsewhere, was not forthcoming until 1975, despite numerous scientists attempting to construct one. Collins also notes that scientists who constructed their attempts and based their efforts on written information alone were all completely unsuccessful in replicating. Only a scientist who had personal contact with the originator of the device was successful in constructing a working TEA laser. The amount of effort and money

involved would greatly exceed the money and effort involved in attempting replications in parapsychology. Moreover the number of failed replications prior to the first success was very large, certainly much greater than the point in parapsychology where, often, the research community decides that some experimental line had 'had it' and proceeds to give up on it.

Pratt (1978) has listed many 'failed' research lines in parapsychology and Millar (1979) has seized upon his claim that since 1882 one could pick almost any date in parapsychology and find some parapsychologist hopefully claiming the 'repeatable experiment' (whatever that is). Even if one could that would prove nothing; a parapsychologist would have to give good reasons why a particular line of experiment should be expected to be repeatable. A claim alone means nothing. In any case many of Pratt's examples are either not repeatable in principle (e.g., work with selected subjects) or else are, actually, quite reasonably repeatable (e.g., the sheep-goat effect). But even if Pratt's listing were realistic, this would not testify to the fact that parapsychology lacks repeatable results; it testifies to the fact that parapsychologists are all too willing, for some reason, to give up on leads.

Finally I want to stress the importance of theoretical systems in parapsychology. The many values of theories in science is too complex a matter to discuss here, but some should be obvious. Moreover we have some eminently testable and predictive models such as the PMIR/Conformance Behaviour models of Stanford (1974a, 1974b, 1977a, 1977b, 1978), the internal attention states model of Honorton (1978) and Braud (1975, 1978) and the Schmidt (1976)/Walker (1975) observational models. Testing these models is an essential priority of research in parapsychology. We cannot afford the luxury of doing experiments we like doing, which have no theoretical rationale, just because we like doing them. I'm personally quite proud of the fact that I have not run any completely novel, original research for around some years now; I have concentrated on lines of research which others have initiated and have tried to explore the leads they observed systematically. Parapsychologists are certainly fiercely independent people, but I think we must concentrate our limited resources in the best areas of experimental research if we want to make progress.

ABSTRACT

The term 'significance' is divided into three types: statistical, practical (magnitude of effect) and theoretical and these different forms of significance discussed with reference to repeatability.

Following dichotomizations of repeatability (within versus between researchers, concrete versus conceptual) problems of assessing repeatability of results are discussed with reference to (1) publication policies (2) box-score analyses and forms thereof and (3) empirical yardsticks.

An appeal is made for theoretical considerations to be paramount in future psi research.

REFERENCES

Note: Because of the large number of references, most titles of articles are omitted.

Barber, T.X. 'Pitfalls in human research', London, Pergamon, 1976.

Beloff, J. 'Presidential address', Research in Parapsychology, 1972.

Braud, L.W. and Braud, W.G. J. A.S.P.R., 1974, 68, 229-245.

Braud, W.G. Journal of Communication, 1975, 25, 142-152.

Braud, W.G. Chapter I of 'Psi and states of awareness', New York, Parapsychology Foundation, 1978.

Braud, W.G. Research in Parapsychology 1978, 1979.

Braud, W.G. and Braud, L.W. J. A.S.P.R., 1974, 67, 26-46.

Broughton, R.S. Parapsychology Review, 1979, 10, no 1, 11-16.

Carpenter, J.C. Chapter in Wolman, B.B. (ed) 'Handbook of Parapsychology', New York, Van Nostrand Reinhold, 1977.

Child, I.L. and Levi, A. J. A.S.P.R., 1979, 73, 273-289.

Child, I.L., Kanthamani, H. and Sweeney, V.M. Research in Parapsychology 1976, 1977, 91-93.

Collins, H.M. Science Studies, 1974, 4, 165-186.

Collins, H.M. Paper presented in symposium at AAAS meeting, Washington, 1978.

Evans, C. Comments in interview for OMNI, November, 1979.

Eysenck, H.J. 'Sense and nonsense in psychology', Harmondsworth, Penguin, 1957.

Gruber, E. Paper presented to 22nd Annual PA Convention, 1979.

Guilford, J.P. 'Fundamental statistics in psychology and education, New York, McGraw-Hill, 1956.

Haraldsson, E. and Johnson, M. E.J.P., 1979, 3, 11-20.

Honorton, C. Presidential Address in 'Research in Parapsychology 1975', 1976.

Honorton, C. Chapter in Wolman, B.B. (ed), op cit, 435-472.

Honorton, C. Chapter in Shapin, B. and Coly, L. (eds) 'Psi and states of awareness', New York, Parapsychology Foundation, 1978 (a).

Honorton, C. Paper presented in symposium at AAAS meeting, Washington, 1978 (b).

Keane, P. and Wells, R. 'Research in Parapsychology 1978', 1979.

Kennedy, J.E. and Taddonio, J.L. J.o.P., 1976, 40, 1-33.

Lane, J. et al Paper presented to Fourth International S.P.R. Conference, Brighton, 1980.

Millar, B. E.J.P., 1979, 3, no 1, 78-110.

Moskovitz, H.R. 'Odors in the environment', in Carterette, E.C. and Friedman, M.P. (eds) 'Handbook of perception' Vol. X, New York, Academic Press, 1978.

Nicol, J.F. and Humphrey, B. J. A.S.P.R., 1953, 47, 133-178.

Nicol, J.F. and Humphrey, B. J. A.S.P.R., 1955, 49, 125-156.

Palmer, J. J. A.S.P.R., 1971, 65, 373-408.

Palmer, J. J. A.S.P.R., 1972, 66, 1-26.

Palmer, J. Chapter in Wolman, B.B. (ed) 'Handbook of parapsychology', New York, Van Nostrand Reinhold, 1977.

Palmer, J., Khamashta, K. and Israelson, K. J. A.S.P.R., 1979, 73, 333-348.

Parker, A. Chapter in Shapin, B. and Colin, L. (eds), op cit, 1978.

Pratt, J.G. J. A.S.P.R., 1978, 72, 127-139.

Ransom, C. J. A.S.P.R., 1971, 65, 289-307.

Sargent, C.L. E.J.P., 1977, 2, no 1, 36-59. (a)

Sargent, C.L. J.o.P., 1977, 41, 275-293. (b)

Sargent, C.L. J.o.P., 1977, 41, 81. (c)

Sargent, C.L. 'Intrasubject and extrasubject factors in psi test performance', unpublished Ph.D. thesis, University of Cambridge, 1979.

Sargent, C.L. J. A.S.P.R., in press (Letter to editor). (a)

Sargent, C.L. 'Exploring psi in the ganzfeld', New York, Parapsychology Foundation, 1980. (b)

Sargent, C.L. in Research in Parapsychology 1979, 1980. (c)

Sargent, C.L. 'Personality and individual differences', 1981, in press.

Sargent, C.L. and Matthews, G. Research Brief 23rd Annual PA Convention 1980.

Schmidt, H. J. A.S.P.R., 1976, 70, 267-291.

Schmitt, M. and Stanford, R.G. J. A.S.P.R., 1978, 72, 177-182.

Sondow, N. J. A.S.P.R., 1980.

Stanford, R.G. J.o.P., 1967, 31, 280-289.

Stanford, R.G. J. A.S.P.R., 1974, 68, 34-57. (a)

Stanford, R.G. J. A.S.P.R., 1974, 68, 321-356. (b)

Stanford, R.G. Presidential Address in 'Research in Parapsychology 1973, 1974. (c)

Stanford, R.G. Chapter 'Experimental Psychokinesis' in Wolman, B.B. (ed), op cit, 1977. (a)

Stanford, R.G. Chapter on 'Conceptual Frameworks' in Wolman, B.B. (ed), op cit, 1977. (b)

Stanford, R.G. J. A.S.P.R., 1978, 72, 197-214.

Stanford, R.G. and Schroeter, W. in 'Research in Parapsychology 1977', 1978.

Stanford, R.G. et al J. A.S.P.R., 1975, 69, 127-133.

Targ, R. and Puthoff, H. Chapter in 'Mind at large', New York, Praeger, 1979.

Tart, C.T. Chapter in Wolman, B.B. (ed), op cit, 1977.

Tart, C.T. Comments made in discussion section of Shapin, B. and Coly, L. (eds), op cit, 1978.

Terry, J. and Honorton, C. J. A.S.P.R., 1976, 70, 207-217.

Ullman, M., Krippner, S. and Vaughan, A. 'Dream telepathy', London, Thurnstone, 1973.

Walker, E.H. Chapter in Oteri, L. (ed) 'Quantum physics and parapsychology', New York, Parapsychology Foundation, 1975.

White, R. Symposium paper presented at 22nd Annual PA Convention, 1979.

Carl L. Sargent
The Psychological Laboratory
University of Cambridge
Cambridge CB2 3EB
United Kingdom