EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 6

NUMBER 1

NOVEMBER 1985

EDITORIAL STAFF

Martin Johnson
Sybo A. Schouten

EDITORIAL CONSULTANTS

E. Bauer, Germany

J. Beloff, England

P. Cassoli, Italy

R. Ejvegaard, Sweden

E.Haraldsson, Iceland

J.M. Houtkooper, The Netherlands

H.H.J. Keil, Australia

H. Kreitler, Israel

S. Otani, Japan

E. Servadio, Italy

R. White, U.S.A.

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

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.

CONTENTS

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.

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

CONTENTS

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.

EUR.J.PARAPSYCHOL., V.6, 1985

PERCEPTUAL DEFENSIVENESS, GANZFELD AND THE PERCIPIENT-ORDER EFFECT

Two experiments

Erlendur Haraldsson and Loftur Reimar Gissurarson
University of Iceland

In several experiments the Defence Mechanism Test (DMT) has been found a relatively successful predictor of ESP performance (Haraldsson, 1978; Haraldsson and Johnson, 1979; Johnson and Haraldsson, 1984; Kragh and Smith, 1970). The DMT-ESP experiments conducted so far have, however, all used forced-choice ESP tasks and none have tested the predictive value of the DMT for free-response ESP performance. Most of the free-response experiments in recent years have used the Ganzfeld technique and many yielded statistically significant results (Honorton, 1985; Sargent, 1980; Schouten, 1981).

The main objective of the present experiment was to test the capacity of the DMT to predict performance of a free-response ESP task in the Ganzfeld. As such, this experiment is a continuation of earlier experiments testing the effect of perceptual-cognitive defensiveness on ESP performance. In working out the details of the experimental design we benefitted from a prepublication manuscript of Hyman's

----

Note: This paper fulfils the publication policy of this journal criticism of the Ganzfeld experiments, and the care that has been taken in describing the procedure can be considered a response to his critical review (Hyman, 1985).

The second objective of this experiment was to test the 'percipient-order effect' hypothesis, which the first author found in an earlier series of experiments (Haraldsson, 1972) and then confirmed in a later experiment (Haraldsson, 1980) but was unable to confirm once more in still another experiment (Houtkooper and Haraldsson, 1985). In all these experiments only vasomotor reactions were used as indicators of ESP by making plethysmographic recording of changes in blood volume in the fingers of the percipients. Emotional stimuli (names emotionally charged for the percipient), which an agent tried to 'send' to a percipient lying in a restful position in a different room, were used as targets. Emotional stimuli are known to cause changes in blood volume in the extremities and can be measured by the plethysmograph.

In the above plethysmographic experiments a pair of subjects was used for each session, one of the pair acting as percipient and the other acting as agent. Immediately after the first session the subjects changed roles. The first percipient then took on the role of the agent and the first agent became the second percipient. The first percipient tended to obtain ESP scores below the mean chance expectation, the second percipient above, with the resulting difference in ESP performance between the first and second percipients being statistically significant. This effect was termed 'percipient-order effect'.

It is of particular interest to test the percipient-order effect with Ganzfeld ESP, as this is somewhat similar to plethysmographic ESP, while there have been reported very few cases of psi-missing with Ganzfeld ESP.

PILOT EXPERIMENT

The pilot experiment was conducted to get acquainted with the Ganzfeld technique, the problems that may be involved in it, and to test the target pool. Since the pilot experiment had the same tight controls and procedures as the experiment that followed, two hypotheses were tested in the pilot experiment.

H1: The Ganzfeld condition will effect extra chance ESP scoring.

H2: The 'percipient-order effect' will operate in the data, such that the scores of the first percipient will be lower than that of the second percipient in each pair that participates in a session.

METHOD

Subjects

Four subjects were members of a psychic development group and four were students. All were believers in psychic phenomena and had some interest in the experiment. A total of 32 Ganzfeld sessions were conducted, with a subject participating in 1 to 7 sessions.

Basic procedure

In every session a pair of subjects participated. First one subject acted as percipient and the other as agent. A second session followed immediately after the first in which the subjects changed their roles, such that the first percipient became agent and the first agent took over the role of the second percipient. This procedure was used to test the occurrence of the 'percipient-order effect'.

When a pair of subjects arrived at the lab, they were welcomed by the two experimenters (EH and LRG). The nature of the Ganzfeld technique was explained to them, they were shown the lab and told that immediately after each session we would examine the results together. A coin was tossed by one of the experimenters to determine which of the subjects would first act as percipient. One of the subjects was asked if he would prefer to act first as percipient or agent. If the emblem turned up in the coin-toss, then that subject obtained the role which he stated a preference for; if the krona turned up, he started in the other role.

From then on the first experimenter (E1, EH) took care of the agent and brought him/her to the agent's room, whereas the second experimenter (E2, LRG) handled the percipient and showed him to his room. Two rooms and two corridors were between the percipient's and agent's rooms (see floor plan in Figure 1). The total distance between the two rooms was about 30 feet. The percipient laid comfortably in a reclining chair ('Lazy Boy'), where E2 placed on him a headphone and two yellowish ping-pong ball halves over his eyes. A 60 watts white lamp was lighted 18 inches from the percipient's face.

When E2 had given instructions to the percipient, he put on the tape-recorder to which the headphone was attached and left the percipient's room for the adjacent room where he stayed till the end of the session. He closed all three doors between the two rooms. The percipient first heard through his headphone light classical music (Mozart) for a few minutes, followed by one minute of silence. Then the percipient heard the pleasant sound of a small river till the end of the session. A microphone in the percipient's room was attached to a tape-recorder in the adjacent room of E2. Through a headphone E2 listened to all statements made by the percipient who had been told to "think out aloud", tell and express every thought or feeling that came to his mind. Whatever he reported was recorded on tape and also written down simultaneously by E2.

E1 took the agent to the agent's room to sit on a comfortable chair. Then he went to an adjacent room (between E2's room and agent's room) to select the target picture for the session.

Target pool

100 pictures were selected by EH as target pool. They were drawn from a large number of pictures from National Geographic Magazine, Iceland Review (which carries beautiful pictures from Iceland), and art prints bought in museums and shops in New York, Amsterdam and Rome. The 100 pictures were in approximately equal numbers drawn from these three sources. All targets were glued on thick 18cm x 26cm cards.

The pool consisted of 25 different fixed target sets, each set containing four pictures that differed widely in content. The pictures in each set were hand shuffled and then numbered from 1 to 4, with the number written on the back of each picture and on the opaque envelope which contained each picture. The four targets (envelopes) of each set were placed in a larger envelope and numbered in the same fashion from 1 to 25. For each set there was an exact copy. The first set was used

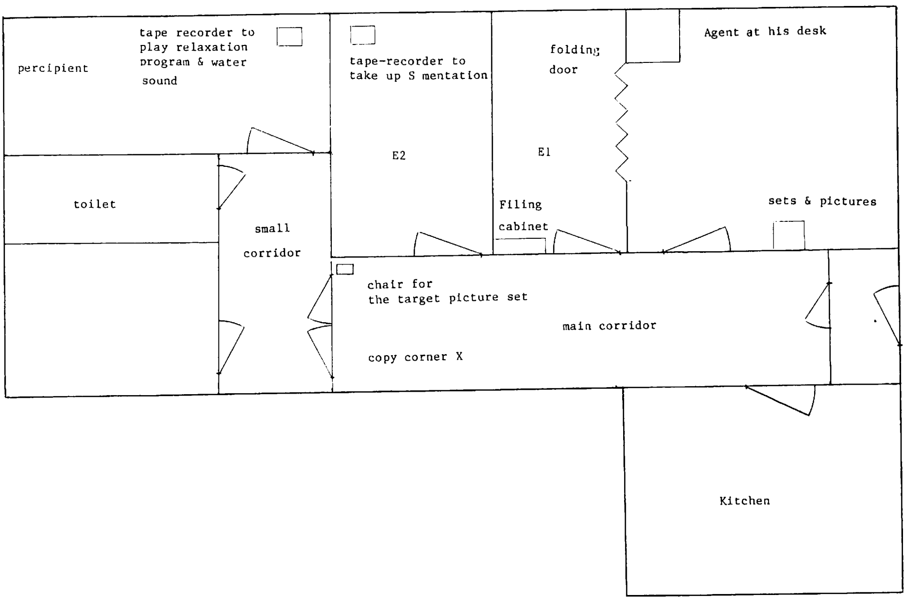

Figure 1
The laboratory

for the agent and the copy set for the percipient's evaluation, which took place immediately after the Ganzfeld session. The two sets of the target pool were kept in separate drawers in a locked filing cabinet in the room adjacent to the agent, with the key remaining with EH (E1).

Randomization of the targets

The randomization was conducted in two stages by E1 while the percipient was being prepared for the experiment by E2. First, the set was randomized and then the picture within the set.

Randomization of the set: 1) A die was thrown 3 times to find the line number (the number at the left margin of each line from 00000 to 19999 in 'A Million Random Digits with 100,000 Normal Deviates', published by the RAND Corporation). Each throw was recorded on a form along with the other information for each session. 2) These three throws of the die made a three digit number, anywhere from 111 to 666, and it determined the line to be used. 3) A fourth throw determined in what block (there are 10 blocks of numbers on each line) the set number would be selected. The first two-digits numbered from 01 to 25 found in that block of the line onwards (continuing with the next line if necessary) determined the number of the set to be used.

Randomization of the pictures within a set: 1) The die was thrown 3 times as before to select a line of random numbers. 2) The die was thrown once more to select a block. 3) The first digit between 1 and 4 in this block of the line or the following ones determined the number of the picture to be used.

The two copies of the selected set were picked out of the filing cabinet. The envelope with the target picture was taken out of the set envelope and taken to the agent. The 'copy' set for the percipient and the rest of the set for the agent were left on the desk in the agent's room. When E2 returned to his room he signaled E1 by knocking on a locked door between E1's and E2's rooms, that the Ganzfeld period had started.

Instructions to percipient

"Keep your eyes open as much as possible throughout the session. First you will hear some beautiful music helping you to relax as well as you can. Then you will hear a sound of a brook. After I have placed the headphones over your ears, I will leave and monitor you from the next room. I will turn out the main ligths before I leave. When the Ganzfeld period is over in about half an hour I will turn them on again".

(Percipient already seated in the reclining chair): "You may release your clothes if they are too tight, your belt, collar and so forth. Sit as comfortably in your chair as you can. Try to be completely relaxed all the time. When you hear the river sound, begin to observe your mental processes and start 'thinking out loud'. Continue to share all your images, thoughts, and feelings, which pass through your mind, with me throughout the session. Try not to concentrate on anything special. Do not bother about, or cling to the sender and his picture, he will do all the concentration work. At some point during the session he will send you a message. Do not try to guess, anticipate or conjure up the message. It will appear in consciousness at the appropriate time. The only thing you have to do is to relax, but try to stay awake all the time. Notice and observe your impressions and thoughts without strain and do not cling to any of them. Just observe them as they go by and speak fluently and continually about everything you think of: ideas, impressions, emotions, sensations, and feelings, sudden impulses, and pictures, etcetera".

After the Ganzfeld situation, in the rating period: "The picture will very seldom fit exactly. You are supposed to evaluate what picture is most likely to correspond to your thoughts, feelings and ideas during the Ganzfeld period. Perhaps there will only be a couple of items in one picture that corresponds to your mentation report".

Instructions to the agent

When E1 had finished the randomization procedure and filled in the form for the session, and E2 had signalled the start of the Ganzfeld period, E1 left his room, closed the door of his room and the agent's room that he now entered. He gave the following instructions to the agent:

"In this envelope is the target picture that we are using in this experiment. We have 100 pictures and I have randomly selected this picture from our pool. The pool is divided into 25 sets, each containing four different pictures. In a little while I will open this envelope and hand you the picture. Look carefully and attentively at the picture, imagine that you are seeing what the picture is actually of, and at the same time think of your companion sitting in the other room. Imagine that he is seeing the picture, that its contents are in his mind and that he is surrounded by it. Try to make the picture a living reality in your mind and hope that the percipient in the other room comes to think of the contents of the picture".

The sending period started after 15 minutes of Ganzfeld (5 minutes of music and 10 minutes of river sound). Then E1 took the target out of the envelope and handed it to the agent who kept it for 20 minutes. The experimental session thus lasted for 35 minutes. The percipient was not told exactly when the sending period would begin.

At the end of the Ganzfeld period E2 went back into the percipient's room and released him of the headphone and the ping-pong balls. E2 read to the percipient the written notes on his mentation report. The percipient was then asked to add any further associations or impressions he may have had from the Ganzfeld session.

When the Ganzfeld period was over and E2 had gone to the percipient's room (E1 heard E2 slam his door), E1 went to E1's room, picked up the copy set and placed it on a chair in the corridor, the arrangement of which can be seen in Figure 1.

E2 fetched the 'copy' set with the target picture, which E1 had placed on the chair in the corridor. Then E2 handed the percipient the 4 pictures of the set and asked him to rank them according to their similarity to his mentation while in Ganzfeld. The percipient was also asked to explain why he ranked the pictures the way he did from the most similar to his report and memory to the least similar. He was also asked to score them on a scale from 0 to 100. These scores were only gathered for eventual post-hoc analysis of how confident the subject was of the similarity between the target and his images or mentation report. These scores were not to be used to test our hypotheses.

Summary of the roles of E1 and E2

The two experimenters took on their separate roles when the experiment was ready to start, the subjects having been shown the rooms and the procedure etc. explained.

First experimenter, E1 (EH):

1) E1 took the agent to his room.

2) E1 went to his (E1's) room and filled out the form for the session.

3) Using a random procedure (involving eight dice throws and 'Million Random Digits') E1 selected the target picture set and then the target picture in the set.

4) E1 took the envelope of the target picture set and the copy set out of the locked filing cabinet. He took the target picture (in envelope) out of the set envelope and went with it to the agent's room closing both doors on the way.

5) E1 gave the agent instructions and chatted with him if time allowed. After 15 minutes of the experimental session E1 handed the agent the target picture, who started 'sending' and did so the next 20 minutes.

6) E1 waited until E2 had gone out of his room at the end of the session and into the percipient's room (the slamming of the doors could easily be heard). E1 then opened his door, picked up the copy set in E1's room and placed it on a chair at the end of the main corridor.

7) E1 walked back into his room, picked up the envelope of the target picture set, returned to the agent's room and took the target picture from the agent and placed it back into its envelope and placed that envelope into the set's envelope. Then he and the agent waited for E2 and the percipient to enter the room.

8) E2 and percipient entered the agent's room after they had finished the evaluation of the set's pictures and recorded the results.

9) E2 placed the pictures on a table in the order that the percipient had ranked them. E1 asked the percipient to describe to him and to the agent why he had ranked each picture the way he had. After the percipient had done that, and perhaps after some questions and discussions, E1 revealed to percipient and E2 what picture of the set was the actual target.

10) If this was the subjects' first session of the day, they changed roles and started another session, the percipient becoming the agent

and the agent becoming the percipient.

11) E1 took the set and copy set envelopes back to the filing cabinet.

Second experimenter, E2 (LRG):

1) When E1 had thrown the dice to determine the first role of each subject, E2 took the percipient to the percipient's room while E1 took the agent to the agent's room.

2) E2 placed the ping-pong halves and the headphone on the percipient who was sitting in the reclining chair.

3) E2 turned on the tape-recorder for the headphones, as well as the 60 watts lamp, which was situated at an 18 inches distance from the face of the percipient. He then left the percipient's room.

4) E2 entered his room (beside the percipient's), closing all three doors on his way. Fifteen minutes after he had turned on the other tape-recorder, he turned on the tape-recorder in his room, which was connected to the percipient's microphone, so that everything the percipient uttered was recorded.

5) E2 wrote down every word the percipient said during the 20 minute sending period.

6) E2 filled out a special form for each session during the experiment.

7) E2 left his room after 35 minutes and went to the percipient and released him of the equipment.

8) E2 asked the percipient to add to his mentation report, any associations, feelings or impressions he may have had, and these were written down by E2 and also tape-recorded.

9) E2 read the mentation report to the percipient.

10) E2 fetched the set with the target picture, which at this time was lying on a chair in the corridor. (Neither E2 nor the percipient at this time knew which of the four pictures had been chosen as the target picture). E2 asked the percipient to evaluate all 4 pictures and rank each one of them, the picture corresponding most with the mentation being given the lowest number (1). The percipient was also asked to tell why he had ranked the pictures in the way he did. He was asked to grade each picture in the set on a scale from 0 to 100.

RESULTS

Direct hits (p=.25) were used as units of analysis. Nine direct hits were obtained in 32 sessions, which is close to mean chance expectation. Thus our hypothesis of extra chance scoring in the Ganzfeld condition was not confirmed. Two female subjects, who both claimed many spontaneous psychic experiences, obtained direct hits that were of rather impressive quality. One of these subjects obtained 2 hits in 5 sessions and the other 3 hits in 6 sessions.

In 16 sessions the first percipients obtained 2 hits, whereas the second percipients obtained 7 hits in an equal number of sessions. Again our hypothesis was not confirmed but approaches significance (chi2=2.47, df=1; transformed into z=1.57, p=.06, one-tailed). In view of the few number of trials involved (32) this result may perhaps be taken as some support for the percipient-order effect (i.e. that second percipients shall obtain more hits than first percipients) since it is so clearly in the expected direction.

EXPERIMENT ONE

The purpose of this experiment was to test two hypotheses:
H1: There is a relationship between ESP in the Ganzfeld condition and perceptual defensiveness as measured by the DMT.
H2: The 'percipient-order effect' will operate in the data, such that the scores of the first percipients will be lower than those of the second percipients.

METHOD

The basic procedure was the same as in the pilot study. The target pool remained the same except for one picture that was replaced because in the pilot experiment it was found to be similar to another picture in the same set.

Twenty unpaid male students from various departments of the University of Iceland each participated in the experiment on two occasions, each time once as percipient and once as agent. Thus we had two Ganzfeld ESP sessions from each subject. Several weeks prior to the Ganzfeld session 48 unselected subjects had been administered the DMT as a part of a different experiment (Haraldsson and Hoeltje, in press). From that group the 15 highest ranking and the 15 lowest ranking DMT scorers were selected as potential subjects. They were paired together irrespective of their DMT scores, which remained unknown to them as well as to E2. Some of them declined to participate or did not show up, but in the end 20 subjects were tested in 38 experimental sessions.

RESULTS

In the 38 Ganzfeld sessions which were conducted, 10 direct hits were obtained, with 28 misses. In the subjects' 20 first Ganzfeld sessions 6 hits were obtained, and 4 in their 18 second sessions. The correlation between the two sessions was only tau=.02.

A calculation of the non-parametric correlation coefficient Kendall's tau between the DMT and the two combined Ganzfeld sessions yielded tau=-.12 (n=20, n.s.). The hypothesis concerning the relationship between perceptual defensiveness and Ganzfeld ESP was thus not confirmed. The correlation between the DMT and the first Ganzfeld sessions was tau=-.31 (n=20, P=.12, two-tailed), and was tau=.21 (n=20, n.s.) in the second sessions. (The two subjects who did not participate in the second session were each given the expected value of .25 for the second session.) Distribution of the DMT scores of the subjects can be seen in Table 1.

The percipient-order effect hypothesis was not confirmed since the first percipients obtained 5 hits, and the second percipients also had 5 hits in an equal number of sessions (chi2=0, df=1, n.s.).

Since all subjects had participated earlier in two ESP computer games, we examined how their Ganzfeld ESP performance related to their forced-choice performance in the computer games. The combined Ganzfeld ESP performance (session 1 and 2 combined) and the combined forced choice ESP performance (computergame 1 and 2) gave a correlation of tau=.20 (n=20, p=.29, two-tailed). Further details can be seen in the correlation matrix in Table 2.

TABLE 1
Means of different ESP tasks in relation to DMT scores

DMT	Ss	No of hits in Ganzfeld Session 1	2	combined	No of hits in Computergame Game 1	2	combined
1	1	1.00	0.00	1.00	11.00	19.00	30.00
2	3	.33	.33	.67	8.00	9.67	17.67
3	3	.67	0.00	.67	7.67	7.00	14.67
4	2	0.00	0.00	0.00	9.50	7.00	16.50
6	3	.33	.33	.67	11.33	9.00	20.33
7	4	0.00	.50	.50	10.75	10.50	21.25
8	2	.50	.13	.63	11.50	8.50	20.00
9	2	0.00	.13	.13	9.50	4.50	14.00
Total	20	.30	.23	.53	9.80	8.90	18.70

DISCUSSION

The relationship between the DMT and Ganzfeld ESP was not significant; in fact the direction was opposite to the one found in experiments with forced-choice ESP. This finding raises some questions. The relationship in the previous Icelandic DMT-ESP experiments has been consistent but weak. In retrospect one may question the soundness of testing our DMT-ESP hypothesis with only 20 subjects. In previous experiments the number of subjects has ranged from 37 to 54, and only two gave independently statistically significant scores, although all gave results in the expected direction. If our hypothesis is correct it would probably be more reasonable to expect results comparable to earlier ones after a further 20-30 sessions, assuming that the relationship is as strong.

Another question concerns the difference in scoring behaviour

TABLE 2
Kendall's tau correlation coefficients between DMT, Ganzfeld and computergames ESP scores
(N=20)

	1	2	3	4	5	6
1 Ganzfeld session 1						
2 Ganzfeld session 2	.02					
3 Ganzfeld combined	.71	.62				
4 ESP Computergame 1	.01	.14	.10			
5 ESP Computergame 2	.33	-.02	.20	.13		
6 ESP Computergames comb.	.30	-.02	.20	.47	.75	
7 Defense Mechanism Test	-.31	.21	-.12	.23	-.03	.01

between forced and free response ESP tasks. The Ganzfeld ESP seems to be more frequently unidirectionally psi-hitting, at least according to Ganzfeld proponents (Honorton, 1978; Sargent, 1980). The forced-choice ESP task is very much a conscious effort, and hence likely to underlie whatever perceptual-cognitive distortions any such task may suffer. A subject in the Ganzfeld situation is not consciously making (or supposed to make) any conscious effort, at least not in our experiment. His or her task is only to register and relate whatever ideas and impressions 'fall into' his mind. One can speculate whether perceptual-cognitive defensiveness is not less likely to operate in such a condition. In psychoanalytic theory free-association (which, in many ways resembles the Ganzfeld) has been hailed as a way to mitigate the operation of defence mechanisms. If such a speculation makes any sense, then the DMT-ESP relationship should be weaker in free-associative ESP tasks than in the forced-choice ones. However, we are on thin ice as we have very little data on which to base a judgement. Only further research can answer the questions that have been raised here.

The percipient-order effect hypothesis was not confirmed but fared somewhat better. The percipient-order effect has been found statistically significant in three of five previous experiments, and the results have been in the same direction in all five. In the present experiments the results approached significance in the pilot experiment, but in experiment one the first and second percipients obtained an equal number of hits. The odd thing about this effect is that the reason for its occurrence is obscure. In a way it resembles a reverse decline effect, so often found in forced-choice experiments. Perhaps it is related to what difference there may be in the nature of free response versus forced-choice ESP tasks. We intend to continue studying it.

The positive scoring in the present experiments was not up to the claims sometimes made for such experiments (Honorton, 1978). It became the impression of the authors that the chance of obtaining significant positive scoring with unselected subjects was meager, but that further testing of subjects who in previous experiments had obtained quantitatively impressive hits might be a promising undertaking. This may be in line with Sargent's and Honorton's observations that successful Ganzfeld experiments have generally been conducted with subjects who have been chosen for the experiment because of successful pretests.

ABSTRACT

The purpose of the present study was to test 1) the relationship between perceptual defensiveness, as measured by the Defence Mechanism Test (DMT), and ESP performance in the Ganzfeld condition, and 2) the occurrence of the percipient-order effect. In one experiment 20 male non-psychology students participated pairwise in two Ganzfeld sessions after being administered the DMT in an earlier session. The DMT-ESP hypothesis was not confirmed. The percipient-order effect was tested in a pilot Ganzfeld ESP experiment in which 8 subjects completed 32 Ganzfeld sessions and in the experiment described above. The results were not significant but in the expected direction in one experiment. When all six experiments in which the percipient-order effect has been tested are combined the overall result is significant.

ACKNOWLEDGEMENTS

A grant for this research from the Icelandic Science Foundation is greatly acknowledged, also helpful suggestions for statistical evaluation by Joop Houtkooper.

REFERENCES

Haraldsson, E. 'Vasomotorische Reaktionen als Indikatoren aussersinnlicher Wahrnehmung', Doctoral dissertation, University of Freiburg, 1972.

Haraldsson, E. 'ESP and the Defense Mechanism Test (DMT): A further validation', E.J.P., 1978, 2, 104-114.

Haraldsson, E 'Confirmation of the percipient-order effect in a plethysmographic study of ESP', J.o.P., 1980, 44(2), 105-124.

Haraldsson, E. and Johnson M.U. 'ESP and the Defense Mechanism Test (DMT). Icelandic study III: A case of experimenter effect?', E.J.P., 1979, 3, 11-20.

Haraldsson, E. and Hoeltje, C. ' The Defense Mechanism Test (DMT) as a predictor of ESP performance. Icelandic study VII.', J.o.P., in press.

Honorton, C. 'Psi and internal attention states: Information retrieval in the Ganzfeld', in: B. Shapin and L. Coly (eds.): Psi and states of awareness, New York, Parapsychology Foundation, 1978.

Honorton, C. 'Meta-analysis of psi Ganzfeld research: A response to Hyman', J.o.P, 1985, 49, 51-91.

Houtkooper, J.M. and Haraldsson, E. 'Experimenter effects in a plethysmographic ESP experiment', E.J.P., 1985, 5(4), 313-326.

Hyman, R. 'The Ganzfeld psi experiment: A critical appraisal', J.o.P., 1985, 49, 3-49.

Johnson, M.U. and Haraldsson, E. 'The Defense Mechanism Test as a predictor of ESP scores', Icelandic studies IV and V', J.o.P., 1984 , 48, 185-200.

Kragh, U. and Smith, G.W. (eds.) ' Percept genetic analysis', Lund, Gleerup, 1970.

Sargent, C.L. 'Exploring psi in the Ganzfeld', Parapsychological Monographs, 1980, 17, New York, Parapsychology Foundation.

Schouten, S.A. 'An overview of details of published Ganzfeld studies', Research Letter, 11, Parapsychology laboratory, University of Utrecht, 1981.

E. Haraldsson
Department of Psychology
University of Iceland
Reykjavik
Iceland

EUR.J.PARAPSYCHOL., V.6, 1985

A FREE-RESPONSE STUDY IN A REAL-LIFE SETTING

Sybo A. Schouten
Jan Merkestein
University of Utrecht

From discussions about the existence of ESP it appears that most authors believe that experimental research will not be able to prove the existence of a psi process unless a repeatable demonstration of psi phenomena can be presented. Of course, what really would constitute a 'repeatable demonstration' seems to be more a question of personal taste than of clear logic (Johnson, 1974). From the fact that spontaneous ESP experiences are relatively rare it can be concluded that, if such a psi process exists, natural conditions tend to inhibit its occurrence. This implies that the assumed psi phenomena are not repeatable by nature, as for instance the daily rise of the sun is. Hence if repeatability is possible at all it can only be forced by creating the proper conditions. But to be able to do so we must first learn what these conditions are. Therefore in our field the repeatable experiment will be the result of our research efforts rather than a condition which must be fulfilled before research is to be considered meaningful. It also follows that experimental ESP research should be process-oriented rather than aimed at proving the existence of an ESP

Note: This paper fulfils the publication policy of this journal

process.

A main obstacle for the progress of process-oriented ESP research is the difficulty in obtaining ESP phenomena under laboratory conditions. Many experimental studies yield no indications of ESP and hence from these studies, despite the sometimes clever design, no conclusions can be drawn. As to the reasons for this lack of results one can only guess. It might well be that psi is a purely spontaneous process by nature and that all efforts to control it are doomed to fail. Another possible reason is lack of knowledge of the proper conditions and of optimal research methods.

A psi experiment can be considered as an attempt to measure variables, among others the variable psi. The research method is the instrument by which psi is measured. But because of our lack of knowledge as to the nature of the psi process it is difficult to devise an instrument which is optimal for measurement. It might well be that our research methods are unsuitable to measure psi and that the current methods are as effective as if we tried to measure auditory signals by means of a voltmeter. Thus it is conceivable that a different research method might yield a higher proportion of successful psi experiments simply because of its higher sensitivity to detect psi. Therefore, it seems worthwhile to explore new approaches in the hope of finding more effective research techniques compared to the current methods. For more than half of this century one research technique has dominated the field: the card-guessing paradigm. Fortunately in the last decades several attempts have been made to apply other research approaches, e.g. animal research, psychophysiological methods, free-response studies, etc.. Not all of these new approaches have been successful, but without these efforts we would not have known which ones. Others, for instance the combination of the free-response technique and Ganzfeld condition, seem more promising (see for instance Honorton, 1985; Hyman, 1985).

Another factor which might be of relevance to the relative lack of progress in process-oriented research is the tendency to focus mainly on analyses which are supposed to yield evidence of ESP and less on analyses which might give insight in less spectacular aspects of the experimental approach. The latter approach involves all kinds of analyses which might yield useful information but which are not part of the main aim of the study. Most ESP experiments are rather complicated and many variables, especially psychological ones, are difficult to control in part for lack of experimental data about them.

That subjects should select the target picture in a free-response study by matching each possible target with their mentation during the session is rather essential to the philosophy behind this approach. But how certain are we that this really happens? Undoubtedly the subject presents us only with a selection of his or her inner experiences, the ones which can be verbalized and which he is able and willing to communicate to us. In addition we know little about how the subject carries out the required matching, if indeed a matching is done at all. Perhaps subjects sometimes rate target pictures mainly on their aesthetic attractiveness.

It is conceivable that a research approach could become more effective and more consistent in its results if we were better able to control variables like the judging procedure. But these are mainly psychological problems, not strictly belonging to the field of parapsychology. One could argue that in many respects progress in process-oriented psi research is strongly linked to progress in traditional psychological research. But often psychology is not able to provide us with the necessary knowledge or research tools and so it seems unavoidable that many of these problems have to be tackled by parapsychologists themselves. However, considering the lack of manpower in the field it is not surprising that researchers appear not to spent too much effort on such problems.

The relative lack of efforts to carry out all sorts of exploratory analyses on the data might also be due to the critics assertion that increasing the number of analyses invalidates the significance of any result observed (see among others Hyman, 1985). This criticism is essentially correct but one can argue that its validity depends on the conclusion one wants to draw from a study. Strict adherence to this criticism might have the effect that the number of analyses in an experiment is strongly reduced and limited to only analyses aimed at testing the main hypothesis in order to avoid possible 'erosion' of the main results by further analyses. However, this strongly reduces the effectiveness of a study in terms of potential information gain versus investment in time and money. A possible solution to this problem, adopted in the present study, could be to distinguish between a section 'Results', in which the analyses related to the main hypotheses are reported and a section 'Further Analyses', in which all kinds of analyses are reported which the author feels to be of interest. Since the 'Results' section involves only analyses which were planned beforehand and which are related to the hypotheses under investigation, the data reported here can be considered to have evidential value. On the other hand, the analyses reported in the section 'Further Analyses' are a selection from a large number of possible analyses and consequently can only be of suggestive value.

The present study was undertaken with two aims in mind. First we wanted to explore a new research method: a free-response study based on daily-life experiences. The second aim was to explore methods for applying a more objective system to select targets based on the subjects' experiences to replace the currently adopted purely 'subjective' judgment of the subject or external judge.

The occurrence of spontaneous paranormal experiences, experiences in daily life to which the experient attributes a paranormal character, suggests that in daily life conditions are occasionally optimal for ESP. We may assume that subjects are able to distinguish between 'normal' or 'explainable' inner experiences, for instance, seeing an acquaintance reminds him of a dinner date with that person, and experiences which are unusual and not easy to explain. The latter might contain some paranormal elements. A first attempt to base an experimental study on such experiences was made by Boerenkamp and Camfferman (1983) who obtained marginally significant results. In the present study the method explored was to run trials lasting a whole day and to ask subjects to note any inner experiences during the day which seemed 'striking' or 'unexplainable'. Then at the end of the day the subject was presented with a set of possible targets and requested to rate them all for degree of agreement with these experiences.

It is known that spontaneous paranormal experiences mainly involve topics which have a strong relevance for the percipient (see among others Schouten, 1979). Therefore it was decided not to apply a neutral target set but to ask each subject to list the five most important topics in their life. These topics were employed as the target set. As a consequence for each subject a different target set was applied. When, as is common in most free-response studies, neutral visual target are employed it can be assumed that the perception of the target picture by agent and percipient will not differ strongly. A picture of a skyscraper will look like a skyscraper to both of them. However, when as in our study targets are employed which are highly personal for the percipient, strong discrepancies might arise between agent's and percipient's mental representation of such targets. For instance, if a percipient rates 'her daughter' as an important topic in her life, her perception and feelings about her daughter will undoubtedly differ from those of the agent who does not know the girl.

The main condition introduced in our study involved this variable. Is it of importance that the agent partly shares the views of the percipient about the target or is it sufficient that the agent knows only the topic and creates her own impression of it?

The second aim of the study concerns the judging procedure. The common procedure in free-response studies is to request subjects to match each possible target with their mentation (in our experiment the 'unexplainable' inner experiences of the day) on degree of agreement. This is a fairly essential element of this research method. However, unfortunately we know little about how such matchings are done. Hence the matching can be considered as a variable which is not under control and which can therefore seriously weaken the effectiveness of free-response studies. If subjects in different conditions of a free-response study apply different matching procedures (which we simply don't know), it might well be that a significant difference observed between conditions is partly or mainly due to this difference in matching. Not surprisingly some authors prefer to have independent and trained judges to do the matching, as for instance Palmer et.al (1977, p.140) who states that "in most cases our money is on a good independent judge".

Matching inner experiences with a set of possible targets involves two main aspects. One is to select from all the experiences and their possible interpretations the ones to use for the matching. The second is to evaluate the degree of correspondence between each experience and the possible targets. Both are rather complex tasks, especially in the case when targets of a complex and multi-detailed nature are employed.

In the present study it was attempted to explore methods which would allow the matching to be carried out in a more objective way. It might well be that some type of experiences are more suited for psi-mediation than others. In that case the matching task could be simplified by selecting only those experiences for the matching. Thus the methods explored involved different ways of selecting specific sets of experiences and to apply only the selected experiences for the matching. If a feasible way of selecting experiences can be found it might become possible to instruct subjects how to select from their own experiences the ones to base the matching on and to have the actual matching done by raters applying an objective system for establishing degree of agreement. For the selection of subsets of experiences the following criteria have been applied: 1) spontaneous, unexpected inner experiences; 2) personal memories; 3) dreams and daydreams 4) experiences related to a topic the subject had forgotten was included in the target set by the time the experiment was carried out; 5) experiences related to the mood of the day.

METHOD AND PROCEDURE

The two authors acted as experimenters. M. was responsible for all dealings with the subjects, whereas S. was responsible for preparing the targets and for supervising the agent. The experimenters did not exchange information about details of their respective tasks until after completion of the study. Thus S. did not know the identity of the subjects who participated in the study and M. did not know the identity of the agent. For communication purposes the subjects' first names were used. Both experimenters were blind to the target and the condition for each trial.

M. found twenty subjects willing to participate in this study. No rewards were offered. Each subject made drawings of the five most important topics at that moment in his or her life and characterized each topic by a word or a short description, for instance: my animals; jogging; worries about my children. These topics were employed as target sets.

S. prepared two target sets for each subject, one made up of the original drawings and one made up of the words characterizing the drawings. Each possible target was put in an opaque envelope on which only the christian name of the subject was printed. A staff member of the laboratory otherwise not connected with the study was then given the two sets of envelopes. He did not know which set contained the drawings and which set contained the words. After shuffling the envelopes of both sets he randomly placed one set on top of the other and gave the by now one set of ten envelopes back to S.. Thus S. only knew that the first and last five envelopes contained the same type of target (drawings or words), but not which type it was nor how the individual targets were distributed over the two sets. S. subsequently numbered the envelopes of the set 1 to 10 respectively.

Random target numbers were provided by another staff member of the laboratory. An Apple RNG board (Bierman board) generated twenty pairs of two random numbers in such a way that if the first number was in the range 1-5, the second number would be in the range 6-10 and vice versa. The first number of the pair indicated the number of the target envelope for the first trial of the subject, and the second number determined the envelope for the second trial. Pairs of numbers were assigned to the subjects by S. based on the order the appointments were made. Thus to the subject for which first an appointment was set up, the first pair of numbers was allocated, etc. This randomization procedure ensured that of the two trials for each subject one would involve a drawing and one a word as target but that the experimenters could not know which condition or target was involved in a trial. It also follows from the procedure that in the two trials of one subject the same topic could be the target.

Only one trial a day was run. M. made appointments for trial days and communicated them by telephone. S. then took the target envelope with the proper target number and ensured that the agent received the envelope for the trial day in time. To reduce the possibility of errors the trial date was written on the envelope.

Throughout the experiment a female 16 years old acted as agent. She was instructed to open the envelope for the trial of the day after she arose in the morning and to carry the target all day with her. She was asked to concentrate occasionally during the day on the target. To ensure her involvement and for checking purposes she was requested to write down in a notebook at the end of the day details about the target issue (topic, drawing or word) and about her impressions associated with that issue. The agent kept all targets until the experiment was completed.

The subject was given an object (an earring) belonging to the agent to carry during the day of the trial. M. instructed the subject to pay attention to any inner experience or unusual event during the day which might be related to the targets. At the end of the day M. visited the subject and recorded the experiences reported by the subject. Then the subject rated each possible target for agreement with the experiences reported on a scale ranging from 0 to 30. For these matchings M. employed a set of photocopies of the drawings provided by the subject. The ratings of the subjects constituted the raw data for the evaluation of the main hypotheses. Subjects were not aware that the two trials were run under different conditions.

Twenty subjects took part in the experiment, 11 males and 9 females. The age range was 16 to 60 years with an average of 28 years.

After the experiment was completed but before the agent was asked to return the target envelopes which would reveal the identity of the targets for each trial M. applied the above mentioned criteria for selecting specific sets of experiences. Thus, when applying the criterion 'personal memory' he singled out from all experiences reported by the subject those which explicitly reflected memories. For instance a statement like: "This afternoon I had to think of the classroom in my high school were Dutch language was taught. I thought about the classroom, not the teacher, a dull room but a positive feeling". Then for each set of criteria he compared for each trial the content of the statements fulfilling the criterion with the possible targets and selected from the target set one target which seemed most strongly related to the statements involved. Hence in these exploratory analyses no ratings were applied, but for each trial just one possible target was selected. Because not in all trials experiences were reported which fulfilled the selection criterion the total number of trials for each selection criterion was less than 40.

It was decided in advance to apply Stanford's Z-score method as described in his 1974 paper (Stanford and Mayer, 1974; Stanford and Sargent, 1983) for evaluating the two main hypotheses of the study.

These hypotheses are: 1) A significant positive deviation will be obtained for the ratings of the target topics over all trials and 2) A difference in ratings for the target pictures will be obtained in the two conditions. The results of selecting targets by the experimenter from subsets of experiences were to be evaluated separately for each subset of experiences by calculating the cumulative probability of the observed number of hits and more extreme number of hits based on the binomial distribution for p=.2 and the total number of matchings involved in the subset.

RESULTS

As a result of the random procedure it turned out that with 11 subjects a drawing was employed as a target in the first trial. Consequently the other 9 subjects started their two trials with a target word. In the case of three subjects the same topic served as target for both sessions. Unfortunately it turned out that with these three subjects a drawing was target in the first session, thus decreasing a possible effect from the conditions introduced.

For each trial the rating of the subject for the target drawing or target word was expressed in a z-score according to the procedure first applied by Stanford. The average z-score of the target topics in the 40 trials is -.217 (sd=.84). Tested against an expected mean of zero the t-test yielded a non-significant T=-1.63 (df=39, p=.11). Hence it can be concluded that no correlation exists between the subjects' ratings of target topics based on their experiences during the day and the actual target topics. Not surprisingly the difference in z-score distributions for target topics between the two conditions appeared to be equally nonsignificant (t=.61, df=19).

Subjects provided on average 5.75 experiences of a different nature which they felt to be possibly related to the target for that day. No relationship is observed between target condition and number of experiences reported. The average number of experiences is 7 for the first trials and 4.5 for the second trials. Fourteen subjects contributed to the lower average number of experiences reported for the second trials by reporting a lower number of experiences for these trials compared to the first trials. An overview of the distribution of conditions and z-scores obtained is presented in table 1.

FURTHER ANALYSES

As described above it was attempted to develop a judging procedure which would enable the experimenter or a judge to apply a more objective procedure for carrying out the matching. Five different methods were applied for selecting subsets of experiences. Some of the selection criteria were selected before the experiment was carried out (spontaneous experiences and memories) and the others were selected by M. based on his experience after recording impressions by subjects. For instance, after a number of trials M. became aware that sometimes subjects' impressions were strongly influenced by the 'mood' of the day. So he decided to try out 'mood of the day' as one of his selection criteria.

Spontaneous unexplainable experiences. In 16 trials subjects reported experiences which they considered strange and unexplainable. Based on these experiences the experimenter selected in 13 trials one possible target and in the remaining three trials two possible targets. It turned out that in 3 of the above mentioned 13 trials and in 2 of the 3 remaining trials the correct target topic was identified. These

TABLE 1
Conditions and results for all subjects

Subject	First trial condition	z	Second trial condition	z
1	d	-1.35	w	0.76
2	w	-1.15	d	-0.65
3	w	-0.46	d	-0.60
4	d	-1.10	w	-0.57
5	d	1.34	w	-1.16
6	w	-0.40	d	-0.78
7	w	-0.41	d	-0.24
8	d	1.29	w	-0.05
9	w	1.16	d	-0.46
10	d	-0.95	w	-0.27
11	d	0.36	w	-0.46
12	d	-0.54	w	-0.46
13	d	-0.45	w	-1.34
14	w	-1.11	d	-0.66
15	w	1.34	d	-0.65
16	w	0.16	d	0.16
17	d	1.86	w	-0.94
18	d	0.52	w	-1.22
19	w	-0.75	d	0.20
20	d	0.18	w	1.18
.S				

Note: d: drawing; w: word

deviations are not sufficiently strong to warrant further interest.

Memories. In 10 trials the subject reported experiences which clearly reflected personal memories. For each of these trials the experimenter selected one possible topic which he considered most strongly related to the memory involved. It appeared that with this procedure the correct topic was identified in 7 out of the 10 trials. According to the binomial distribution the probability of obtaining under these conditions 7 or more hits is p=.00008. The target condition (drawing or word) appears not to correlate with the outcome of this analysis. In 4 of the 'hits' a drawing and in 3 of the 'hits' a word was employed as target. The 10 trials in which memories were reported appear to be equally distributed over the drawing and word target conditions. No relationship was found between the number of memories and total number of experiences reported over the subjects.

Dreams. Only in four trials experiences were reported which appeared to be related to dreams or dream-like states. In no instance did the topic selected based on these experiences match the target for that trial.

The forgotten target topic. In the course of the experiment it was observed that in four trials the subject did not remember one of the topics of the target set. When these forgotten topics are compared with the real target issues for those trials no agreement was observed. Hence the assumption that perhaps some defense-mechanism would make the subject to forget the topic which was the target for the day, a possible form of psi-missing, was not substantiated.

Mood and main activities of the day. Equally unsuccessful was the attempt to find the real targets by selecting the topics which best matched the main mood or activity of the day of the trial. This analysis involved 8 trials of which none led to a correct identification of the target topic.

DISCUSSION

As regards its main hypotheses this experiment did not yield the results hoped for. The combination of a free-response technique with subjects rating possible targets on the basis of daily-life experiences did not result in significant scores. Nor did the incorporation of subject-related target sets made up of issues important to the subject enhance the scores. Hence the present approach does not seem a promising one.

The only suggestive finding appears to be that high scoring was obtained when the experimenter acted as judge and selected targets based on only experiences reflecting personal memories. As discussed in the introduction this finding can only be of suggestive and not of evidential value. Although one would rather expect spontaneous unexplainable experiences to be related to the target, the notion that ESP impressions might be mediated by memories is not a novel one. Indeed, as can be concluded from an overview article by Roll on 'ESP and memory' (in White, R.A., 1976) the notion is nearly as old as the field itself. But mainly studies have been carried out investigating the correlation between properties of memory and ESP (see for example, Irwin, 1979). However, the relationship between memories and content of ESP impressions appears never to have been explored systematically. Only occasional references to the subject can be found. In spontaneous cases it is sometimes observed that the ESP impressions are mediated by memories. A practical problem in this type of research is, however, that from the reports it is often not possible to decide whether the ESP impressions were based on or associated with a memory. In free-response studies also occasionally references to this subject are found. In Honorton and Harper's (1974) publication on their first Ganzfeld experiment one of the first observations they report after discussing the results is that "memory, in particular, seems to serve as a vehicle for psi-mediation." Sondow, who is one of the few authors who focussed in her Ganzfeld study on the problems associated with the judging remarked: "It has often been noticed that psi sometimes manifests in the form of reactivated memories - and, as can be seen in some of the illustrative samples shown -, qualitative material in the present experiment supports the observation" (Sondow, 1979, p.142). Probably more of such observations can be found scattered throughout the literature.

On the other hand, the relationship can not be so strong to the extent that nearly all psi impressions are mediated by memories. In that case undoubtedly more experimenters would have noticed and reported it. If we may attach any value to the relationship between targets and memories reported above, it might be that the relationship emerged so strongly in this study because highly personal targets were employed. At any rate the result seems sufficiently encouraging to investigate further the possibility of finding a more objective way for selecting ESP targets in free-response studies.

ABSTRACT

A free-response study has been carried out in which subjects based

the ratings of targets on unexplainable or noteworthy experiences during the day of the trial. Twenty subjects participated for two trials each. Each subject provided a target set to be used only for that subject consisting of the five most important topics in their lives at that time. Targets were presented to the agent either in the way the subject had provided it or as a concept so that the agent had to form her own representation of it. In order to find a more objective way to carry out the matching after completion of the trials but before the targets were known, one of the experimenters tried to select targets himself by applying different criteria for selecting experiences to base the matching on.

No significant result was observed as regards the total number of hits based on subjects' matchings or as regards the difference between the two conditions of agent's representation of the targets. One attempt to select targets based on objective criteria resulted in 7 hits in 10 trials. Since this result is a selection from a number of analyses it can only be considered to be of suggestive value.

LITERATURE

Boerenkamp, H.G. and Camfferman, G. 'Applying daily experiences in experimental ESP research', E.J.P., 1983, 4, 435-455.

Honorton,C. and Harper, S. 'Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input', J. A.S.P.R., 1974, 68, 156-168.

Honorton, C. 'Meta-analysis of psi Ganzfeld research: A response to Hyman', J.o.P., 1985, 49, 51-93.

Hyman, R. 'The Ganzfeld psi experiment: A critical appraisal', J.o.P., 1985, 49, 3-51.

Irwin, H.J. 'On directional inconsistency in the correlation between ESP and memory', J.o.P., 1979, 43, 31-40.

Johnson, M.U. 'Models of control and control of bias', Inaugural Speech, State University of Utrecht, November 1974.

Palmer, J., Bogart, D.N., Jones, S.M. and Tart, C.T. 'Scoring patterns in an ESP Ganzfeld experiment', J. A.S.P.R., 1977, 71, 121-145.

Roll, W.G. 'ESP and memory', In White, R.A. (ed.): Surveys in Parapsychology, Scarecrow Press, Metuchen, N.J., 1976, 351-382.

Schouten, S.A. 'Analysis of spontaneous cases as reported in Phantasms of the Living', E.J.P., 1979, 2, 408-454.

Sondow, N. 'Effects of associations and feedback on psi in the Ganzfeld:Is there more than meet the judge's eye', J. A.S.P.R., 73, 1979, 123-149.

Stanford, R.G. and Mayer, B. 'Relaxation as a psi-conducive state: A replication and exploration of parameters', J. A.S.P.R., 1974, 68, 182-191.

Stanford, R.G. and Sargent, C.L. 'Z-scores in free-response methodology: Comments on their utility and correction of an error', J. A.S.P.R., 77, 1983, 319-327.

EUR.J.PARAPSYCHOL., V.6, 1985

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART III. THE FIRST GROUP OF EXPERIMENTAL SERIES

Hendrik G. Boerenkamp
University of Utrecht

In a previous paper the design of an experimental study of paranormal impressions of psychics was described (Boerenkamp, 1985a). Psychics are persons who believe themselves able to obtain paranormal impressions at will. Usually, psychics are consulted by clients about problems related to themselves and to persons in their environment. The complete study consists of a predetermined number of series of sessions with a group of psychics and non-psychics in which the subjects gave their impressions about persons unknown to them. The series with the psychics were divided in three standard series, plus two subgroups of five experimental series each. Two of the three standard series were held at the start of the investigation (standard series A and B). These standard series were included to provide a description of the content of sessions when psychics apply their assumed psi abilities under conditions which resemble as much as possible the daily circumstances of sessions with clients (see Boerenkamp, 1985b, for the results). A further aim of including these series was to compare the content of these sessions with the content of the sessions of each of the experimental series, which were held to study the effect of a number of variables on the content of the sessions, i.e. the statements of the psychics about target persons. The experimental series were divided in two subgroups of five series each. The second subgroup was carried out when the first subgroup was completed and was introduced to the psychics as being more

'experimental' compared to the series in the first subgroup.
In this paper the results of the first subgroup of experimental series are discussed and compared with the results of the two standard series.

In the standard series the psychic was presented a photograph and an object of a target person. Target persons were chosen from the environment of the sitter. The psychic was then invited to give paranormal impressions concerning this person. The psychic was aware that the sitter was acquainted with the target person, but he or she was not informed beforehand about special problems in the life of the target person. The psychic received immediate feedback to his statements in the form of an affirmation or a denial (confined informative action), occasionally followed by some clarification (extended informative action), which provided additional related information. No information unrelated to the topic being discussed was provided in the feedback.

The first subgroup of five experimental series was included to study the effect of the following variables:

(1) The type of inductor in feedback conditions (series 1 and 2).
In the standard series the psychic was given an object as well as a photograph of a target person to be used as the inductor. In series 1 the psychic was only given an object belonging to the target person, e.g. a ring or a pipe. In series 2 the psychic was given a photograph of the same target person. The psychic was not informed that in the two sessions the same person served as target person. The session of series 1 always preceded the session of series 2. In the session of series 2 the sitter reacted with only confined informative actions when topics were discussed about which extended information was provided for that target person in the session of series 1.

(2) The amount of feedback (series 3 and 4).
In the standard series the sitter reacted with confined (about 60%) as well as extended informative actions (about 40%). In series 3 the sitter reacted with confined informative actions only. In series 4 the sitter reacted with no feedback at all, i.e. the sitter reacted only with the uninformative actions "hmm-hmm" and "I see" (in order to make clear that he was listening to the statements).

(3) The importance of the problem of the target person (series 5).
In the standard series the consultations concerned various persons

with different problems which were not of acute importance. In series 5 the psychics were consulted about a problem which was of acute importance, i.e. a missing person. In this series before the start of the session the psychics were informed about the nature of the problem.

The data from the sessions are described in terms of the number of statements by the psychic and the number of informative actions (feedback) by the sitter. Also the results of the informational, structural and interactional analyses are presented. In the informational analysis it is established which of the psychic's statements are sufficiently specific and spontaneous to assign them potential paranormal value, taking into account the available information at the moment the statement was made and whether those potential paranormal statements are correct. Such statements are called statements with positive paranormal value. In the structural analysis various characteristics of the set of statements are studied, e.g. the topics discussed in the statements. The interactional analysis concerns the actions psychics take when the sitter denies the correctness of the content of a statement, e.g. giving another interpretation of the impression. For a more detailed discussion of these concepts, see Boerenkamp, 1985a.

The two standard series A and B were carried out with 12 psychics. The experimental series 1,2,3 and 4 were conducted with 11 of the 12. In series 1 and 2 the contribution of P8 is lacking because this psychic died before the second session was held. In series 3 and 4 the contribution of P6 was lost because of a malfunctioning of the tape-recorder in one of the sessions. Series 5 was carried out with 14 psychics as two psychics (P13 and P14) were included who specialized in cases of missing persons.
From the previous study (Boerenkamp, 1985b) it appeared that the behaviour of the psychics in standard series A was very consistent with their behaviour in standard series B. Hence, the data of standard series A and B were first combined in order to minimize the role of accidental influences on the content of the standard series. The combined data of the two standard series are compared with the data of each of the experimental series. (The combined data of standard series A and B are further denoted as standard series A+B). The comparison between the data of standard series A+B and the data of each of the experimental series is only based on the data of the psychics involved in both series. Thus the data of P8 were excluded from the standard series A+B in the comparison between standard series A+B and series 1

and 2, the data of P6 were excluded from the standard series A+B in the comparison between standard series A+B and series 3 and 4, and the data of P13 and P14 were excluded from series 5 in the comparison between standard series A+B and series 5. Moreover, the statements of P13 and P14 were used only in the informational analysis. As they appeared to be of no special importance (a very low number of statements with positive paranormal value), they are left out in the further analyses.

Number of statements and informative actions

The number of statements made by the psychics in the different series is presented in table 1.

TABLE 1
Number of statements by the psychics

	A+B	1	2	3	4	5	total
P1	295	64	95	104	54	104	716
P2	129	41	66	40	26	39	341
P3	219	84	68	90	42	74	577
P4	147	42	37	29	14	107	376
P5	197	66	35	27	52	51	428
P6	213	78	111	--	--	84	486
P7	184	70	70	66	45	120	555
P8	81	--	--	38	21	32	172
P9	134	64	57	46	25	38	364
P10	259	165	87	126	68	86	791
P11	137	95	59	42	33	52	418
P12	126	29	42	61	35	72	365
all	2121	798	727	669	415	859	5589

In series 1 the psychics made 798 statements; the number of informative actions by the sitter was 562 (210 extended and 352 confined ones). In series 2 the psychics made 727 statements; the number of informative actions by the sitter was 567 (197 extended and 380 confined ones). Thus, it appeared that the psychics made about the same number of statements in series 1 (object of the target person as inductor) and series 2 (photograph of the same target person as inductor). Further it appeared that the number of informative actions by the sitter was nearly equal in series 1 and 2 and that the relative number of extended reactions in the two series was not significantly different (chi-square= 2.19, df= 1, n.s.). (In this paper all chi-square tests applied are two-tailed). Thus presenting a photograph or an object of the same target person did not influence the type of feedback.

In the standard series A+B the psychics (minus P8) made 2040 statements while the number of informative actions by the sitter was 1387 (556 extended and 831 confined ones). In series 1+2 the same 11 psychics made 1525 statements, while the number of informative actions was 1129 (397 extended and 732 confined ones). Thus, it appeared that psychics made significantly more statements in the two standard series (2040), in which they used two inductors, than they did in series 1 and 2 (1525), in which they used only one inductor. These results suggest that each of the two inductors applied in the standard series evoked a number of statements. However, the behaviour of the sitters might also have contributed to this difference. Although they gave relatively the same number of feedback responses (1387 responses to 2040 statements in standard series A+B and 1129 responses to 1525 statements in series 1+2; chi-square= 2.55, df= 1, n.s.), they gave it relatively more often in the form of extended feedback in the standard series A+B (40% extended informative actions) than in series 1+2 (35% extended informative actions); (chi-square= 6.20, df= 1, $p< .02$). This means that the sitters had a tendency to give less extended feedback along with giving one instead of two inductors. Therefore, the difference between the number of statements in the standard series A+B and series 1+2 has to be attributed to a combined effect of presenting two inductors versus one and giving extended feedback more often versus less often.

As was the case in the standard series , the number of informative actions by the sitter depended on the number of statements made by the psychic in series 1 and 2. The sitter reacted with more informative actions when the psychic made more statements (series 1: Spearman r= .86, t= 5.01, df= 9, $p< .001$; series 2: Spearman r= .99, t= 44.70, df= 9, $p< .001$). (All correlations applied in this paper are Spearman r

correlations and are denoted further with rs; all t-tests applied in this paper are two-tailed).
Further, it appeared that psychics were consistent in the number of statements they made in the two standard series A and B. Each of them had a consistent preference for the duration of a session. The psychics were also rather consistent in the number of statements in the series 1 and 2 ($rs = .62$, $t = 2.46$, $df=9$, $p < .05$). The correlation between the length of the sessions of each psychic in standard series A+B (minus P8) on the one hand and series 1+2 on the other hand is significant ($rs = .75$, $t = 3.32$, $df = 9$, $p < .01$). This implies that the type of inductor or number of inductors hardly influences the individual preference of the psychics for the duration of a session.

In series 3 the psychics made 669 statements; the number of informative actions by the sitter was 438 (only confined ones). In series 4 the psychics made 415 statements and the sitter made no informative actions. In the standard series the psychics (minus P6) made 1908 statements; the number of informative actions by the sitter was 1308 (509 extended and 799 confined ones). The sitters reacted in series 3 as often as they did in the standard series (chi-square= 0.37, $df = 1$, n.s.). Thus, it appeared that retaining extended feedback reduced the number of statements considerably (669 in series 3 and a number of 954 in a comparable standard series). That the number of statements in a session depended first of all on the amount of feedback appeared from the much lower number of statements in series 4 in which no feedback at all was provided (415 in series 4 and a number of 954 in a comparable standard series).
As was the case in the standard series A+B the number of informative actions by the sitter in series 3 depended on the number of statements made by the psychics ($rs = .86$, $t = 5.14$, $df = 9$, $p < .001$).
Further, it appeared that psychics were rather consistent in the number of statements they made in series 3 and 4 (chi-square= .64, $t = 2.47$, $df = 9$, $p < .05$). The correlation between the number of statements in standard series A+B on the one hand, and in series 3+4 on the other hand, is significant ($rs = .74$, $t = 3.27$, $df = 9$, $p < .01$). This implies that the amount of feedback hardly influences their individual preference for the duration of a session.

In series 5 (important problem) the psychics (minus P13 and P14) made 859 statements; the number of informative actions by the sitter was 529 (187 extended and 342 confined ones). In standard series A+B the 12 psychics made 2121 statements; the number of informative actions by the sitter was 1441 (568 extended and 873 confined ones).

Thus, it appeared that the sitters reacted in series 5 as often as in the standard series (chi-square= 2.19, df= 1, n.s.) and that they did it relatively the same number of times in an extended way (chi-square= 2.54, df= 1, n.s.). Hence a serious problem of which they are aware before the start of the session reduces rather than augments the number of statements (859 statements in series 5 and a number of 1060 in a comparable standard series).
As was the case in standard series A+B the number of informative actions by the sitter depended on the number of statements by the psychic in series 5 (r_s= .73, t= 3.42, df= 10, p< .01).
It also appeared that psychics were rather consistent in the number of statements made in standard series A+B and in series 5 (r_s= .64, t= 2.61, df= 10, p< .04). This implies that the importance of the problem of the target person hardly influences their individual preference for the duration of a session.

The informational analysis

The procedure of selecting statements with potential paranomal value is described in detail by Boerenkamp (1984). The selection was based on the estimate of the probability of correspondence (specific versus vague) combined with the estimate of the degree of spontaneity (spontaneous versus inferred) of each statement by two judges. Thus, each statement was rated by two judges in two different ways on a four-point scale. On these scales a low degree of potential paranormal value is represented by a score of 1 or 2 and a high degree of potential paranormal value by a score of 3 or 4. In the study mentioned above it was found that two judges agreed on 82% of the statements to which of the categories (1 or 2 versus 3 or 4) it should be assigned.

The inter-raterreliability for two judges in the standard series A+B and in each of the five experimental series was comparable to the inter-raterreliability obtained in the previous study. In the standard series A+B the judges agreed on 84% of the statements, in series 1 on 81%, in series 2 on 80%, in series 3 on 82%, in series 4 on 79% and in series 5 on 82% of the statements. The distributions of the scores of potential paranormal value, based on the combined scores of two judges rating each statement on both probability of correspondence and degree of spontaneity on scale ranges from 1 to 4, are presented in table 2.

TABLE 2
Distribution of scores of potential paranormal value in the standard series A+B and the first five experimental series

score	low 4-7	(%)	low-med 7-9	(%)	medium 9-11	(%)	med-hig 11-13	(%)	high 13-16	(%)	total
series A+B	1174	(55)	479	(23)	280	(13)	119	(6)	69	(3)	2121
series 1	426	(53)	164	(21)	129	(16)	49	(6)	30	(4)	798
series 2	359	(49)	170	(23)	111	(15)	55	(8)	32	(4)	727
series 3	385	(58)	144	(21)	94	(14)	26	(4)	20	(3)	669
series 4	187	(45)	109	(26)	79	(19)	23	(6)	17	(5)	415
series 5	424	(49)	282	(33)	100	(12)	38	(4)	15	(2)	859
all series	2955	(53)	1348	(24)	793	(14)	310	(6)	183	(3)	5589

According to the Kolmogorov-Smirnov two-sample test the distributions of scores of potential paranormal value in series 1 and 2 do not significantly differ (Dmax= .040, chi-square= 2.44, df=2, n.s.). (The Kolmogorov-Smirnov two-sample test is denoted further in this paper as KS). The distributions of scores of potential paranormal value in series 1 and series 2 do not significantly differ from the distribution of scores of potential paranormal value in the standard series A+B (minus P8).
The distributions of scores in series 3 and 4 significantly differ (KS: Dmax= .126, chi-square= 16.19, df=2, p<.001). The distribution of scores in series 3 is not significantly different from the distribution of scores in the standard series A+B, but the distribution of scores in series 4 is significantly different from the distribution of scores in the standard series A+B (KS: Dmax= .112, chi-square= 14.59, df=2, p< .001). The statements in series 4 in which no feedback was provided were rated as having higher potential paranormal value than the statements in feedback conditions. However, the difference between the distribution in series 4 and the distribution in standard series A+B is mainly attributable to the lower categories and not to the relevant medium-high and high

categories.
The distribution of scores in series 5 is different from the distribution of scores in the standard series A+B to a marginally significant degree (KS: Dmax= .060, chi-square= 6.76, df=2, $p < .03$). The statements in series 5 were rated as having higher potential paranormal value. However, here again the difference is not mainly attributable to the relevant medium-high and high categories. Applying a cut-off criterion between the medium and medium-high categories, each series yielded about 10% (range 6% to 12%) statements with potential paranormal value.

Of the 188 statements of standard series A+B having potential paranormal value, 105 were made in the first half and 83 were made in the second half of the session. This difference was not significant. Comparable results were observed in each of the five experimental series. In each series more statements with potential paranormal value were made in the first half of the session. However, except for series 2 (chi-square= 7.45, df=1, $p < .01$) the difference within each series was not significant. The preponderance of statements with potential paranormal value in the first half of the session is strongly significant when all series (series A+B and 1-5) are combined (288 in the first half and 205 in the second half; chi-square= 14.93, df=1, $p < .001$).

It further appeared that the psychics were rather consistent in the percentage of statements with potential paranormal value over the different series. The percentages of statements with potential paranormal value of each psychic for the different series are presented in table 3.

The correlation between the proportion of statements with potential paranormal value over psychics in series 1 and 2 is positive to a marginally significant degree (rs= .70, t= 2.98, df= 9, $p < .02$). The correlation between the proportions in series A+B on the one hand and series 1+2 on the other hand is significant (rs= .79, t= 3.90, df= 9, $p < .005$). The correlation between the proportions in series 3 and 4 is positive but not significant (rs= .53, t= 1.89, df= 9, n.s.). The same applies for the correlation between the proportions of statements with potential paranormal value in standard series A+B and each of the series 3 and 4 (series A+B and series 3: rs= .54, t= 1.92, df= 9, n.s.; series A+B and series 4: rs= .50, t= 1.75, df= 9, n.s.). The correlation between standard series A+B and series 5 is positive to a significant degree (rs= .80, t= 4.23, df= 10, $p < .002$). These results

TABLE 3
Proportions of statements with potential paranormal value

	A+B	1	2	3	4	5	total
P1	13%	6%	16%	9%	6%	13%	11%
P2	5%	5%	5%	13%	8%	8%	6%
P3	4%	11%	7%	6%	17%	4%	6%
P4	4%	7%	5%	0%	0%	1%	3%
P5	5%	11%	11%	11%	2%	8%	7%
P6	9%	14%	14%	---	---	7%	11%
P7	21%	11%	16%	8%	27%	8%	15%
P8	11%	---	---	5%	0%	6%	8%
P9	3%	2%	5%	4%	0%	0%	3%
P10	12%	13%	20%	6%	9%	9%	11%
P11	4%	6%	8%	5%	6%	2%	5%
P12	11%	24%	17%	10%	20%	6%	12%
all	9%	10%	12%	7%	10%	6%	9%

suggest that as regards the proportion of statements with potential paranormal value, the psychics are hardly influenced by type of inductor (object versus photograph), number of inductors, and importance of the problem. Only the amount of feedback appeared to have a certain influence. Some psychics take more 'risks' by giving comparatively more spontaneous and specific statements in a situation in which they receive less or no feedback than in a situation in which they receive extended feedback. With other psychics an opposite effect is noticable. However it has to be noted that despite the nonsignificance, the correlations between series 3 and 4 and between standard series A+B and series 3 and 4, respectively, are all positive. This suggests that as regards the proportion of statements with potential paranormal value the impact of the feedback is also restricted.

The proportion of statements with potential paranormal value summarized over the series varied over psychics from 3% to 15%. Hence between psychics there is some variation in this respect and consequently one could make a distinction between more and less risk-taking psychics.

In Boerenkamp (1984) the statements of series 5 were used in order to study whether judges could discriminate sufficiently reliably between statements with potential paranormal value and other statements. In that study two psychics (P13 and P14), who only participated in series 5, were included. Both psychics restricted their psychic activities to cases of missing persons, and both made a relatively high number of statements with potential paranormal value. One of the conclusions in that study was that mediums made significantly fewer statements with potential paranormal value than did the other psychics (chi-square= 7.97, df=1, $p < .005$). However, if the two specialized psychics are excluded from the analysis, there is no significant difference in the proportion of statements with potential paranormal value between mediums and the other psychics in series 5 (chi-square= 2.93, df=1, n.s.). The same applies for all other series.

The second step in the informational analysis was to establish how many of the about 10% of statements having potential paranormal value also fulfilled the criterion of 'sufficient degree of correspondence'. Each statement having potential paranormal value was rated on a four point scale as being true or untrue by the experimenter who was acquainted with the target person. Statements which received a score of either 1 or 2 were called statements of negative paranormal value (untrue statements) and statements which received a score of 3 or 4 were called statements of positive paranormal value (true statements). From table 4 it appears that in all series the degree of correspondence between the content of the statements and the facts about the person is rather low.

No difference was observed between the distributions of positive and negative paranormal statements in series 1 and 2 (KS: Dmax= 0.02, chi-square= 0.06, df= 2, n.s.), nor between the distributions in series 3 and 4 (KS: Dmax= 0.06, chi-square= 0.29, df=2, n.s.). It further appeared that none of the distributions in the experimental series 1,2,3,4 and 5 differed significantly from the distribution in standard series A+B. (Distributions corrected for P8, P6 and P13+P14, respectively).
Table 5 presents for each psychic and for all series combined the percentage of statements with potential paranormal value compared with

TABLE 4
Distribution of scores of 'degree of correspondence'

score	negative value				positive value			
	1	(%)	2	(%)	3	(%)	4	(%)
series A+B	101	(54)	60	(32)	23	(12)	4	(2)
series 1	39	(49)	25	(32)	12	(15)	3	(4)
series 2	44	(50)	25	(29)	13	(15)	5	(6)
series 3	24	(52)	13	(28)	4	(9)	5	(11)
series 4	21	(52)	12	(30)	5	(13)	2	(5)
series 5	31	(58)	16	(30)	4	(8)	2	(4)
all series	260	(53)	151	(31)	61	(12)	21	(4)

all statements (ppv/all), the percentage of statements with positive paranormal value compared with statements with potential paranormal value (+pv/ppv) and the percentage of statements with positive paranormal value compared with all statements (+pv/all).

TABLE 5
Statements with positive paranormal value

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	all
ppv/all	11%	6%	6%	3%	7%	11%	15%	8%	3%	11%	5%	12%	9%
+pv/ppv	13%	31%	24%	8%	4%	8%	16%	2%	10%	15%	24%	29%	17%
+pv/all	2%	2%	2%	0%	0%	1%	2%	2%	0%	2%	1%	3%	1%

It appears that no significant correlation exists between the proportion of statements with potential paranormal value (ppv/all) and the proportion of statements having positive paranormal value (+pv/ppv), (r_s= 0.11, t= 0.35, df=10, n.s.). In other words, spontaneous and specific statements of psychics who most often make such statements have the same probability of being correct as those of the less risk taking psychics. Furthermore, although some variation among psychics exists as regards the proportion of statements with potential paranormal value, there is hardly variation among psychics as regards the proportion of statements with positive paranormal value (spontaneous, specific and correct statements). Hence one can distinguish between more and less 'risk-taking' psychics, but not between more and less 'good' psychics.

The structural analysis

In each of the experimental series the number of statements with positive paranormal value was too low to render meaningful a statistical comparison between distributions of statements with positive paranormal value and statements with negative paranormal value as regards the different characteristics studied. Therefore, for each characteristic the number of statements with positive paranormal value and the number of statements with negative paranormal value were first summarized over all series (series A+B and series 1-5) and then compared with respect to each characteristic.

(1) Topics discussed in the statements

The total number of topics discussed in the 2121 statements of standard series A+B was 2856. The distribution of the topics in the standard series A+B and in each of the experimental series is presented in table 6. The distributions of all statements (all) and statements having potential paranormal value (ppv) are presented in real numbers for the main categories (A-E), whereas for the subcategories (10-54) percentages are given. (For some categories the description is abbreviated because of the limited space in the table).

The distributions of the topics were rather similar for the different series. All correlations between the frequencies of the topics (subcategories 10-45) for series 1 versus series 2, series 3 versus series 4 and standard series A+B versus each of the experimental series) are to a significant degree positive.

TABLE 6
Distribution of topics discussed in the statements

	A+B	1	2	3	4	5
Total number of topics	2856	1138	1014	931	607	1435
A Physical characteristics: all	462	344	275	175	136	126
Physical characteristics: ppv	73	40	35	18	16	12
10 Sex	0%	1%	1%	0%	0%	0%
11 Age	1%	3%	1%	1%	3%	2%
12 Appearance	5%	9%	7%	6%	5%	4%
13 Bodily health	8%	12%	14%	7%	9%	3%
14 Being alive or dead	2%	5%	4%	5%	5%	0%
B Psychological charact.: all	1224	378	373	354	234	380
Psychological charact.: ppv	45	25	25	6	8	15
21 Personality traits	27%	19%	20%	27%	22%	12%
22 Psychological circumstances	14%	13%	16%	9%	15%	13%
23 Religious orientation	2%	1%	1%	2%	2%	1%
C Relations: all	521	143	126	150	70	147
Relations: ppv	26	12	13	7	6	10
31 Relations with family members	8%	7%	5%	6%	7%	7%
32 Relations with friends	7%	3%	3%	7%	4%	2%
33 Relation with sitter	3%	3%	4%	3%	1%	1%
D Specific topics: all	649	273	240	252	167	257
Specific topics: ppv	93	38	48	24	24	22
41 Civil status	3%	3%	3%	6%	5%	2%
42 Circumstances in work	9%	8%	9%	9%	8%	6%
43 Circumstances in living	3%	4%	4%	2%	4%	3%
44 Leisure activity	4%	3%	4%	3%	3%	1%
45 Specific name, event	4%	6%	4%	7%	7%	6%
E Disappearance: all						525
Disappearance: ppv						23
51 Being alive or dead						5%
52 Causes of disappearance						12%
53 Events after disappearance						15%
54 Developments after session						5%

(Distributions corrected for P8, P6 and P13+P14 respectively; categories 51-54 in series 5 excluded). The highest correlation was observed between series 1 and series 2 (r_s= .95, t= 11.00, df= 14, $p<$ .001) and the lowest between series A+B and series 4 (r_s= .79, t= 4.81, df= 14, $p<$ .001).
Since in the sessions of series 1 and of series 2 the same target person was involved, it is of interest to study whether the content of the statements about that target person was consistent over the two sessions. It has to be remembered that the session of series 1 (object) always preceded in time the session of series 2 (photograph) and that in the session of series 2 the sitter gave no extended feedback about topics about which extended feedback was provided in the session of series 1. It appeared impossible to classify the two sets of statements of each psychic as consistent or inconsistent with each other. Although each psychic discussed the different topics in each session about the same number of times, e.g. 2% of the topics about religious and social orientation in each session, the content of the topic most often was different. For example, the psychic says about the target person in the session of series 1 "He is catholic" and in the session of series 2 "He is conservative". If only the two sets of statements with potential paranormal value of each psychic were compared neither a consistent nor an inconsistent pattern appeared. However, while hardly any identical statement having potential paranormal value was made in the two series, there were some clear inconsistencies. For example, a psychic said about the target person in the session of series 1 "She died about five years ago" and in the session of series 2 "She is ill and will die within a year from now". It has further to be noted that none of the psychics noticed that he was talking about the same person a second time.

In standard series A+B the distribution of topics (main categories A-D) discussed in the statements with potential paranormal value (ppv) differred strongly from the distribution of topics discussed in the statements without potential paranormal value (chi-square= 107.05, df= 3, $p<$ .001). In the set of statements with potential paranormal value fewer descriptions of psychological characteristics and relations with other people were observed. The same pattern was observed in each of the experimental series.
As stated above the distributions of topics discussed in the statements with positive paranormal value and in the statements with negative paranormal value were first summarized over all series before being compared. Table 7 presents the distributions of topics discussed in the statements with positive paranormal value (+pv) and negative

paranormal value (-pv) in each of the series. The proportion of statements with positive paranormal value in each of the main categories is nearly equal (chi-square= 3.28, df=3, n.s.). This means that, for example, a statement about the physical health of the target person has the same probability of being correct as a statement about the profession of the target person.

TABLE 7
Distribution of topics for statements with positive and for statements with negative paranormal value

	A Physical		B Psychological		C Relational		D Specific	
	+pv	-pv	+pv	-pv	+pv	-pv	+pv	-pv
series A+B	12	61	6	39	4	22	13	80
series 1	4	36	6	19	4	8	7	31
series 2	8	27	4	21	4	9	10	38
series 3	0	18	0	6	1	6	8	16
series 4	2	14	2	6	1	5	6	18
series 5	0	12	3	12	2	8	2	20
all series	26	168	21	103	16	58	46	203

(2) Person discussed in the topics

Out of the 1138 topics discussed in the statements of series 1, a total of 130 (11%) concerned a person related to the target person. For series 2 those figures were 75 out of 1014 topics (7%). The difference between series 1 and 2 is significant (chi-square= 9.62, df=1, p< .002). In standard series A+B (minus P8) 355 of the 2757 topics (13%) concerned a person related to the target person. The distributions in standard series A+B and series 1 do not significantly differ (chi-square= 1.17, df= 1, n.s.). The distributions in standard series A+B and series 2 are significantly different (chi-square= 20.61, df=1, p< .0001). Thus, an object, for example a ring, induces more impressions about persons related to the target person than a

photograph.
Out of the 931 topics discussed in the statements of series 3 a total of 113 (12%) concerned a person related to the target person. In series 4 the comparable figures were 89 out of 607 topics (15%). The difference between series 3 and 4 is not significant (chi-square= 1.84, df= 1, n.s.). In standard series A+B (minus P6) 233 out of 2562 topics (9%) concerned a person related to the target person. Standard series A+B and series 3 differ significantly in this respect (chi-square= 6.75, df= 1, $p < .01$). The same applies to standard series A+B and series 4 (chi-square= 16.06, df= 1, $P < .0001$). Thus, the number of topics about a person related to the target person increases when the amount of feedback is lower.
Out of the 1435 topics discussed in the statements of series 5, a total of 123 (9%) concerned persons related to the target person. The comparable data in standard series A+B were 357 out of 2856 topics (13%). Thus, the number of topics about persons related to the target person is lower in the case in which the problem of the target person is of acute importance and indicated beforehand (chi-square= 14.44, df= 1, $p < .0001$).

Further, it appeared in standard series A+B that statements with potential paranormal value involved relatively more often topics concerning a person related to the target person (26%) than statements with no potential paranormal value (11%). Apparently when psychics talk about persons related to the target person they are for example less likely to concentrate on the psychological characteristics of these persons and more likely to discuss their appearance. The same pattern was observed in series 1 and 2. In series 1 the set of statements with potential paranormal value contained 20% of topics concerning persons related to the target person but the set of statements without potential paranormal value had only 10% (chi-square= 8.38, df=1, $p < .004$). For series 2 these numbers were 20% and 6%, respectively (chi-square= 15.25, df=1, $p < .0001$). No difference in this respect was observed in series 3 in which the comparable numbers were 16% and 12% (chi-square= 0.60, df=1, n.s.) and in series 4 in which the numbers were 15% and 15%. Thus, the potential paranormal value of statements about the target person is equal to the potential paranormal value of statements about persons related to the target person in confined and no feedback conditions. In series 5 in which the comparable numbers were 21% and 8% respectively, the pattern was in accordance with the pattern in the standard series (chi-square= 14.81, df=1, $p < .0001$).
Finally, it appeared that the set of statements with positive

paranormal value, summarized over all series, had the same distribution with respect to this characteristic as the set of statements with negative paranormal value (chi-square= 0.43, df= 1, n.s.).

(3) Number of statements about past, present and future

For the different series the distributions of statements about past, present and future are presented in table 8.

TABLE 8
Distribution of statements about past, present and future

	past	past %	present	present %	future	future %
series A+B	339	16%	1584	75%	198	9%
series 1	99	12%	619	78%	80	10%
series 2	95	13%	566	78%	66	9%
series 3	116	17%	487	73%	66	10%
series 4	90	22%	259	62%	66	16%
series 5	115	13%	676	79%	68	8%
all series	854	15%	4191	75%	544	10%

The distributions of statements about past, present and future in series 1 and 2 are not significantly different (chi-square= 0.49, df= 2, n.s.). Neither do the distributions in series A+B (minus P8) and series 1 (chi-square= 5.24, df= 2, n.s.), and the distributions in series A+B (minus P8) and series 2 (chi-square= 4.87, df= 2, n.s.) differ.
The distributions in series 3 and 4 are significantly different (chi-square= 14.23, df= 2, p< .001). The distributions in series A+B (minus P6) and series 3 are not significantly different (chi-square= 0.83, df= 2, n.s.), while the distributions in series A+B (minus P6) and series 4 differ (chi-square= 27.25, df= 2, p< .0001). Series 4 (no feedback) involved significantly more statements about the past and

the future, as compared to series 3 (confined feedback), as well as compared to standard series A+B (confined and extended feedback). The distributions in series A+B and series 5 are not significantly different (chi-square= 5.38, df= 2, n.s.).
Therefore, it seems that only the absence of feedback affects the behaviour of the psychics in this respect, resulting in more statements about the past as well as the future.

When one splits up the sessions in first and second halves in order to see whether the session as a whole has a temporal pattern, it appeared that there was a preponderance of statements about the future in the second half of the sessions in the standard series A+B. The distributions of statements about past, present and future for the different series are presented in table 9.

TABLE 9
Distribution of statements about past, present and future in the first and second halves of the sessions

	past		present		future	
	FH	SH	FH	SH	FH	SH
series A+B	162	177	828	756	71	127
series 1	51	48	324	295	24	56
series 2	57	38	292	274	15	51
series 3	47	69	260	227	28	38
series 4	37	53	147	112	24	42
series 5	60	55	359	317	11	57
all series	414	440	2210	1981	173	371

Note: FH: first half; SH: second half.

It appeared that the preponderance of statements about the future in the second half of the sessions was consistent over the different series (chi-square values varying from 7.92, df= 2, p< .02 in series 3 to 33.94, df= 2, p< .0001 in series 5). In addition to the preponderance of statements about the future, in series 3 and 4 also statements about the past preponderate in the second half of the sessions. Hence confined or no feedback conditions evoke the tendency to make statements about past and future towards the end of the session.

Further, in standard series A+B it appeared that the set of statements with potential paranormal value contained relatively more statements about the past (29%) and relatively fewer statements about the present and the future compared to the set of statements without potential paranormal value (15%). This difference was also consistent over the different experimental series (series 1: 22% versus 12%; series 2: 23% versus 13%: series 3: 24% versus 17%; series 4: 33% versus 22%; series 5: 21% versus 13%).
The set of statements with positive paranormal value, summarized over all series, had the same distribution with respect to this characteristic as the set of statements with negative paranormal value (chi-square= 4.68, df= 2, n.s.).

(4) Number of statements about a favourable, neutral or unfavourable state of affairs.

In the comparison between the two standard series A and B the only strong inconsistency between the two series appeared to be the difference in the number of statements about favourable and unfavourable states of affairs concerning the target person and related persons (Boerenkamp, 1985b). This number was influenced to a certain degree by the different ways the psychics perceived the target persons. For instance, in certain cases psychic 1 had a positive impression and psychic 2 a negative impression of the same person. Therefore, interpretions of differences in this respect between standard series A+B and each of the experimental series are tentative, since a part of the difference has to be attributed to this 'standard' variation. The distributions of statements about a favourable, neutral or unfavourable state of affairs are presented in table 10.

The distributions in series 1 and 2 are different to a marginally significant degree (chi-square= 7.77, df=2, p< .03). As will be remembered in series 1 and 2 the psychics made statements about the same target person (in series 1 using an object and in series 2 employing a photograph as the inductor). Whereas the difference between the two standard series A and B primarily consisted of a

TABLE 10
Distribution of statements concerning a favourable, neutral or unfavourable state of affairs

	favourable	fav %	neutral	neu %	unfavourable	unf %
series A+B	419	20%	753	35%	949	45%
series 1	110	14%	318	40%	370	46%
series 2	120	17%	241	33%	366	50%
series 3	170	25%	251	38%	248	37%
series 4	77	19%	168	40%	170	41%
series 5	41	5%	456	53%	362	42%
all series	937	17%	2187	39%	2465	44%

higher number of statements about a favourable state of affairs compared to statements about an unfavourable state of affairs, the difference between series 1 and 2 appeared primarily to be the result of a difference in number of statements about a neutral state of affairs. More statements about a favourable as well as more statements about an unfavourable state of affairs are made by the psychics when they employ a photograph of the person as an inductor. Thus, especially a photograph elicits statements of an 'emotional' nature.
In congruence with this tentative conclusion is the fact that the distribution in standard series A+B (minus P8) is significantly different from the distribution in series 1 (chi-square= 9.33, df=2, $p< .01$) but not significantly different from the distribution in series 2 (chi-square= 3.77, df=2, n.s.).
The distributions in series 3 and 4 are different to a marginally significant degree (chi-square= 6.87, df=2, $p< .04$). The difference is primarily due to a higher number of statements about a favourable state of affairs and a lower number of statements about an unfavourable state of affairs in series 3. The distribution in standard series A+B (minus P6) differs significantly from the distribution in series 3 (chi-square= 20.32, df=2, $p< .0001$) but not

from the distribution in series 4 (chi-square= 5.43, df= 2, n.s.). In addition to the 'standard' variation there are two other conceivable reasons for the difference observed. One is the possibility that because of the confined feedback, which implied that problems of the target person indicated by the psychic were not discussed by the sitter, the psychics were made to believe that the target person had fewer problems. The other reason might be that in series 3 the psychics made more statements about favourable states of affairs in order to elicit more, i.e. extended, feedback from the sitter.
The distributions in standard series A+B and series 5 are significantly different (chi-square= 143.01, df= 2, $p < .0001$). This is not surprising, because it is to be expected that in the case of a missing person the number of statements about a favourable state of affairs is relatively low.

When one splits up the sessions of standard series A+B into first and second halves in order to see whether the sessions as a whole had an 'emotional' pattern, it appeared that statements about an unfavourable state of affairs slightly preponderated in the first half of the session (chi-square= 5.89, df=2, $p < .05$). The distributions of statements in first and second halves of the sessions for the different series are presented in table 11.

It appeared that the slight preponderance of statements about an unfavourable state of affairs in the first half of the sessions was consistent over the series only when considered in absolute numbers. However, in none of the series 1-4 does the difference reach significance. Only in series 5 was a marginally significant difference observed (chi-square= 6.85, df=2, $p < .05$). However, summarized over all series the difference becomes significant (chi-square= 10.29, df= 2, $p < .01$). Possibly this indicates that psychics are accustomed to sitters who indicate problems at the beginning of a session.

The set of statements with potential paranormal value in standard series A+B had the same distribution with respect to this characteristic as the set of statements without potential paranormal value (chi-square= 2.09, df= 2, n.s.). The same holds for each of the experimental series.
The set of statements with positive paranormal value, summarized over all series, also does not differ in this respect from the set of statements with negative paranormal value (chi-square= 1.53, df= 2, n.s.).

TABLE 11
Distribution of statements concerning a favourable, neutral or unfavourable state of affairs in the first and second halves of the sessions

	favourable		neutral		unfavourable	
	FH	SH	FH	SH	FH	SH
series A+B	198	221	361	392	502	447
series 1	54	56	159	159	186	184
series 2	49	71	126	115	189	177
series 3	83	87	127	124	125	123
series 4	32	45	86	82	90	80
series 5	20	21	210	246	200	162
all series	436	501	1069	1118	1292	1173

Note: FH: first half; SH: second half.

(5) Number of statements in the form of advice

In standard series A+B, 120 statements (6%) involved advice. In series 1, this number was 49 statements (6%) and in series 2, 54 statements (8%). The difference between series 1 and 2 is not significant (chi-square= 1.00, df= 1, n.s.). The number of statements involving advice in standard series A+B (minus P8) did not significantly differ from the number of statements involving advice in series 1 (chi-square= 0.44, df= 1, n.s.) nor from the number of statements involving advice in series 2 (chi-square= 3.55, df= 1, n.s.).

In series 3, 36 statements (5%) involved advice and in series 4, 27 statements (7%). The difference between series 3 and series 4 is not significant (chi-square= 0.40, df= 1, n.s.). The number of statements involving advice in standard series A+B (minus P6) was not significantly different from the number of statements involving advice in series 3 (chi-square= 0.12, df= 1, n.s.) and in series 4 (chi-square= 1.40, df= 1, n.s.). Thus, neither the type of inductor

nor the amount of feedback influenced the proportion of statements involving advice in a session.
In series 5 only 5 statements (1%) involved advice. This number is significantly low compared to the amount of advice in standard series A+B (chi-square= 37.94, df= 1, p< .0001). Contrary to expectation the number of statements involving advice turned out to be significantly lower in the case of a missing person compared to less dramatic cases. However, it has to be noted that the difference is based only on statements in which an explicit advice is given. A statement about where a missing target person was to be found, e.g. "the person has drowned in the port of Rotterdam, and is hooked onto a boat", might be interpreted as containing an implicit advice (go and look there). Such statements have not been scored as an advice unless the psychic stated it explicitly as an advice, e.g. "the family should contact the police of Rotterdam". As observed above, the number of statements stating an explicit advice was considerably lower in the case of the more dramatic problem which was the topic of series 5.

In all series the number of statements with potential paranormal value in which advice is given is too low to enable a statistical comparison with the number of statements involving advice in the set of statements without potential paranormal value. Therefore, in this case the analysis is based on all statements summarized over the series. It appeared that the statements with potential paranormal value contained fewer statements involving advice compared to the other statements (3% versus 5%). However, this difference is not significant (chi-square= 3.06, df= 2, n.s.).
Since in all series the number of statements with potential paranormal value involving advice was 17, of which only 1 could be assigned positive paranormal value, it is impossible to apply a chi-square test in order to see whether statements involving advice would be more or less often correct than other statements. However, there is no indication that there would be a difference.

(6) Number of times a silence precedes a statement

In standard series A+B, 337 statements (16%) were preceded by a silence of 3 seconds or more. In series 1, 90 statements (11%) and in series 2, 77 statements (11%). The difference between series 1 and 2 is not significant (chi-square= 0.06, df= 1, n.s.). The number of statements preceded by a silence in standard series A+B (minus P8) was significantly higher than the number of such statements in series 1 (chi-square= 9.03, df= 1, p< .01) and in series 2 (chi-square= 10.51, df =1, $p < .01$). A probable explanation is that in the standard series the psychics shifted their attention within a session several times from the photograph to the object and vice versa, causing more pauses. In series 3, 115 statements (17%) were preceded by a silence and in series 4, 107 statements (26%). This difference is significant (chi-square= 11.09, df= 1, $p < .001$). The number of statements preceded by a silence in standard series A+B (minus P6) is not significantly different from the number of such statements in series 3 (chi-square= 0.74, df= 1, n.s.) but significantly lower than in series 4 (chi-square= 23.47, df= 1, $p < .0001$). Thus, the psychics are more apt to consider what to say next when feedback is completely retained. In series 5, 185 statements (22%) were preceded by a silence. This number is significantly higher than the number in the standard series A+B (chi-square= 13.11, df= 1, $p < .001$). Thus, the psychics are also more inclined to consider more carefully what to say next when the problem is of greater importance.

Further, it appeared that with respect to this characteristic the set of statements with potential paranormal value in standard series A+B is not significantly different from the set of statements without potential paranormal value. However, in series 1 (chi-square= 15.75, df= 1, $p < .001$) and series 4 (chi-square= 12.20, df= 1, $p < .001$), the sets are significantly different. In the two series the set of statements with potential paranormal value contained relatively more statements preceded by a silence. Thus, in the condition in which only an object of the target person was presented or in the condition in which they received no feedback at all, the psychics try to obtain more often spontaneous and specific impressions during the silences. Finally, it appeared that the set of statements with positive paranormal value, summarized over all series, has the same distribution with respect to this characteristic as the set of statements with negative paranormal value (chi-square= 0.04, df= 1, n.s.).

(7) Number of positive and rhetorical statements

Rhetorical statements are statements in which the psychic asks for immediate feedback in contrast which positive statements in which the psychic does not ask for immediate feedback. The distributions of positive and rhetorical statements for the different series are presented in table 12.

The distributions of positive and rhetorical statements in series 1

TABLE 12
Distribution of positive and rhetorical statements

	positive	positive %	rhetorical	rhetorical %
series A+B	1474	69%	647	31%
series 1	559	70%	239	30%
series 2	505	69%	222	31%
series 3	426	64%	243	36%
series 4	356	86%	59	14%
series 5	617	72%	242	28%
all series	3937	70%	1652	30%

and 2 are not significantly different. Neither are the distributions in series A+B (minus P8) and series 1, or in series A+B and series 2. It has to be remembered that in series 3 the psychic was informed that the sitter was acquainted with the target person and in series 4 that the sitter was not acquainted with the target person. Therefore, one would expect only positive statements in series 4. However, it appeared that psychics are so accustomed to receiving feedback that a considerable number of times (14%) they simply neglected the information from the sitter that he was not acquainted with the target person by making rhetorical statements. It appeared that all psychics had this tendency. The percentage of rhetorical statements in series 4 varied between the psychics from 5% to 28% with a median of 14%.
The distributions in standard series A+B (minus P6) and series 3 are different to a marginally significant degree (chi-square= 5.37, df= 1, $p < .03$). In series 3 the proportion of rhetorical statements is higher. Thus, it looks as if the psychics try to elicit more feedback in the condition in which the sitter gave confined feedback only.
The distributions of positive and rhetorical statements in standard series A+B and series 5 are not significantly different (chi-square= 1.48, df= 1, n.s.).

When the sessions of standard series A+B were split up into first and second halves, a significant difference appeared between the first and second halves of the sessions as regards the number of rhetorical statements. More rhetorical statements were found in the first half of the sessions (chi-square= 14.32, df= 1, p< .001). In table 13 the distributions of rhetorical and positive statements are presented for all series.

TABLE 13
Distribution of positive and rhetorical statements in the first and second halves of the sessions

	positi e statements		rhetorical statements	
	FH	SH	FH	SH
series A+B	700	780	361	280
series 1	260	299	139	100
series 2	231	274	133	89
series 3	205	221	130	113
series 4	166	190	42	17
series 5	283	334	147	95
all series	1845	2098	952	694

Note: FH: first half; SH: second half.

The preponderance of rhetorical statements in the first half of the sessions was found to be consistent over the series. Only in series 3 was a non significant difference between number of positive and rhetorical statements in the two parts of the session observed (chi-square= 1.58, df= 1, n.s.). Hence, in series 3 (confined feedback only) the psychics tried to elicit more feedback until the end of the session.

The set of statements with potential paranormal value in standard series A+B contained more rhetorical statements compared to the set of statements with no potential paranormal value (chi-square= 112.24, df= 1, p< .0001). Table 14 presents the distributions of positive and rhetorical statements with potential paranormal value (ppv) and with no potential paranormal value (nppv) for all series.

TABLE 14
Distribution of positive and rhetorical statements with and with no potential paranormal value

	positive statements		rhetorical statements	
	ppv	nppv	ppv	nppv
series A+B	67	1413	121	520
series 1	36	523	43	196
series 2	46	459	41	181
series 3	9	417	37	206
series 4	25	331	15	44
series 5	29	588	24	218
all series	212	3731	281	1365

This finding appeared to be very consistent over the experimental series (chi-square values varying from 7.30, df= 1, $p< .01$ in series 5 to 39.53, df= 1, $p< .0001$ in series 3).
Finally, the set of statements with positive paranormal value summarized over all series did not differ from the set of statements with negative paranormal value (chi-square= 0.41, df= 1, n.s.) as regards number of positive and rhetorical statements.

(8) Number of statements preceding an informative action

Since all sorts of statements might invite feedback it was calculated how many statements were made by the psychic before the sitter reacted with an informative action. If the psychic made a statement and the sitter reacted directly with one or more informative actions, it is indicated by P1S in table 15. If the psychic made two statements before the sitter reacted, it is indicated by P2S etc..

TABLE 15
Distribution of statements preceding an informative action

	P1S%	P1S	P2S	P3S	P4S	P>4S
series A+B	45%	964	273	84	31	39
series 1	49%	390	114	24	9	11
series 2	61%	445	75	20	11	6
series 3	49%	326	59	28	14	13
series 4	---	---	--	--	--	--
series 5	37%	315	78	51	18	21
all series	44%	2440	599	207	83	90

The distributions in series 1 and 2 are significantly different (chi-square= 13.63, df=4, p< .01). The sitter appeared to have reacted more often directly (more P1S interactions) in the sessions of series 2. It has to be remembered that the psychics in series 1 and 2 made statements about the same person, that in time the session of series 1 always preceded the session of series 2 and that the sitter did not give extended feedback on topics which were discussed in an extended way in the session of series 1. Thus, the sitter's awareness of talking a second time about the same target person apparently also resulted in more direct feedback reactions. In congruence with this finding is the fact that the distributions in series A+B (minus P8) and series 1 do not differ (chi-square= 3.77, df= 4, n.s.), while the distributions in series A+B (minus P8) and series 2 are different (chi-square= 24.69, df= 4, p< .0001).
The distributions in series A+B (minus P 6)(extended plus confined feedback) and series 3 (confined feedback only) are not significantly different (chi-square= 8.12, df= 4, n.s.) which implies that giving less feedback did not result in more direct feedback reactions.
The distributions in standard series A+B and series 5 are significantly different (chi-square= 19.10, df= 4, p< .001). This means that the sitter reacted with more reserve (with fewer P1S interactions) in the case of the missing person than he did in the

standard series.

In standard series A+B the set of statements with potential paranormal value contained on the average the same number of statements which were part of P1S interactions as the set of statements with no potential paranormal value (chi-square= 0.02, df= 1, n.s.). A similar result was found in each of the experimental series.
The set of statements with positive paranormal value and the set of statements with negative paranormal value, summarized over all series, did not differ in this respect either (chi-square= 0.01, df=1, n.s.).

The interactional analysis

The subject of this analysis is the actions taken by the psychics after receiving a denial as a reaction to one of their statements. For the standard series it was found that the psychics mainly use four types of responses to a denial:

(1) Accepting the denial by giving another impression (AD)
(2) Giving a new interpretation to the denied impression (NI)
(3) Suggesting the target person knows better (T>S)
(4) Suggesting the content of the informative action is equal to the content of the statement (I=S)

The AD and NI reactions are called the 'acceptance of failure' reactions and the T>S and I=S reactions are called the 'resistance against failure' reactions. In table 16 the different responses of the psychics for the different series are presented.

As was the case in the standard series the psychics felt themselves rather 'safe' with the sitters in the experimental series. They probably experienced a permissive atmosphere for making 'mistakes' as they most often reacted with 'acceptance of failure' reactions in all series.
The distributions in series 1 and 2 are not significantly different (chi-square= 7.19, df= 3, n.s.). The distribution in series A+B (minus P8) is different from the distribution in series 1 to a marginally significant degree (chi-square= 9.60, df= 3, p< .03) but not different from the distribution in series 2 (chi-square= 7.22, df= 3, n.s.). Thus, psychics tend to react somewhat more often with 'resistance against failure' reactions than with 'acceptance of failure' reactions in the condition in which they only use an object as the inductor.

TABLE 16
Distribution of types of responses to a denial

	AD	NI	T>S	I=S	acceptance AD+NI %	resistance T>S + I=S %
series A+B	22	101	25	16	75%	25%
series 1	18	53	15	24	64%	36%
series 2	14	65	6	15	79%	21%
series 3	16	46	13	12	71%	29%
series 4	--	--	--	--	---	---
series 5	19	34	9	14	70%	30%
all series	89	299	68	81	72%	28%

The distributions in standard series A+B (minus P6) and series 3 are not significantly different (chi-square= 3.07, df= 3, n.s.). This implies that the amount of feedback by the sitter did not influence the type of action taken by the psychic after receiving a denial. The distribution in standard series A+B and series 5 are different to a marginally significant degree (chi-square= 10.24, df= 3, p< .02). In the case of a missing person the psychics tend to react more often with accepting the denial (AD) and less often with giving a new interpretation (NI).

The relatively low number of denials in the standard series did not permit an analysis in terms of variation in type of response between psychics in the standard series only. Therefore, in table 17 the different responses of each psychic in standard series A+B and in series 1,2,3 and 5 are presented. In the last column the total number of statements with potential paranormal value in those series is also presented.

It is apparent that all psychics most often reacted with an 'acceptance of failure' reaction (range 54% to 89%, with a median of 73%) and that all psychics did this most often by giving a new

TABLE 17
Distribution of types of responses to a denial among psychics

	AD	NI	T>S	I=S	accept % AD + NI	resist % T>S + I=S	total	total ppv
P1	13	45	10	8	73%	27%	76	82
P2	2	11	4	7	54%	46%	24	21
P3	8	25	8	11	63%	37%	52	37
P4	6	12	1	3	82%	18%	22	12
P5	1	7	1	0	89%	11%	9	28
P6	8	25	2	6	80%	20%	41	52
P7	23	52	6	12	81%	19%	93	82
P8	1	8	0	6	60%	40%	15	13
P9	1	26	14	4	60%	40%	45	10
P10	11	39	12	9	70%	30%	71	90
P11	7	27	8	11	64%	36%	53	21
P12	8	22	2	4	87%	13%	36	45
all	89	299	68	81	72%	28%	537	493

interpretation to the impression (NI). All psychics used all types of responses at least once (with only two exceptions). The correlation between the total number of denials and the total number of statements with potential paranormal value made is positive to a marginally signifcant degree ($r_s = 0.61$, $t = 2.45$, $df = 10$, $p < .04$). In other words, the psychic who took most risks (made most spontaneous and specific statements) received the highest number of denials. However, one might expect that the correlation would be stronger than it is found to be. This relatively low correlation is caused mainly by two psychics who had a somewhat specific style in the sessions. One psychic (P5) tended to talk very much about psychological and relational characteristics. Therefore, one would expect a low number of statements with potential paranormal value as well as a low number of denials. Indeed, she received the lowest number of denials compared with the other psychics. However, in contrast to all other psychics she had a tendency to make a relatively high number of statements in terms of 'hidden' personality characteristics, such as masochistic, sadistic, homosexual etc.. In general these statements could not be commented on by the sitter (after all, they were 'hidden' characteristics) but were rated by the judges as having potential paranormal value. The other psychic (P9) made the most unspecific statements of all the psychics. Therefore, again one would expect a low number of statements with potential paranormal value as well as a low number of denials. Indeed, among all the psychics she made the lowest number of statements having potential paranormal value. However, she had a tendency to make statements relatively often about physical characteristics, especially bodily health, which, although being unspecific, had to be denied often by the sitter as being not true.

DISCUSSION

As regards the length of the sessions, it appears first of all that the number of statements from the psychics in a session depends on the amount of feedback provided by the sitter. In the condition in which the sitter provides maximal feedback (extended as well as confined informative actions), the psychics make more than twice as many statements than they do in the condition in which the sitter provides no feedback. In the condition of intermediate feedback (confined informative actions only), the number of statements is also intermediate. In addition to feedback two other factors appear to have some influence.

The first one is the number of inductors presented to the psychic. When two inductors (photograph and object) are presented, each of them evokes a number of statements. This is also illustrated by the finding that in series 1 and 2 in which only one inductor was given, the psychics made significantly fewer statements preceded by a silence compared to the standard series in which two inductors of the target person were presented. In the standard series within a session they shift their attention several times from the photograph to the object and vice versa, producing more periods of silence. From the equal number of statements in series 1 (object) and series 2 (photograph), it can be concluded that the type of inductor has no significant impact on the number of statements in feedback conditions.

The second factor is the importance of a problem indicated at the beginning of a session. In the case of an acutely important problem psychics generally make fewer statements.

Further, it appears that as regards the number of statements, psychics are rather consistent over the different series. The different experimental manipulations did not have much effect on their individual preference for sessions of a certain length.

In the standard series it was found that approximately only one out of ten statements could be considered as meeting the criterion of 'being sufficiently spontaneous and specific', i.e. could be assigned potential paranormal value and that approximately only one out of ten of these statements with potential paranormal value could be considered to meet the criterion of 'being sufficiently correct', i.e. could be assigned positive paranormal value. Probably none of the statements which met both criteria could be considered to have a probability lower than 1 in 100. Therefore, it was concluded that the observed 'unexplainable' correspondences between the statements and the facts concerning the target person could be satisfactorily explained by assuming chance coincidence and that it is not necessary to assume a specific paranormal ability of the psychics.
The data in each of the experimental series support this tentative conclusion. Although some minor differences were observed in the distributions of scores of potential paranormal value as well as in the distributions of scores of positive paranormal value, there is no indication that any of the variables in the first five experimental series, i.e. type of inductor in feedback conditions, amount of feedback and acute importance of the problem of the target person, had a facilitating or inhibiting impact on the assumed ability.
As stated in the design, definite conclusions about the ability of psychics have to be based on all sessions of the investigation. However, as neither the amount of feedback nor the importance of the problem had any impact on the assumed ability, it is very doubtful whether research with psychics offers a more promising research method than the methods currently employed in parapsychology.
Another important finding in the informational analysis is that the different psychics are rather consistent over the different series as regards the proportion of statements with potential paranormal value in a session.
Finally, it appears that some variation among psychics exists in regard to the proportion of statements with potential paranormal value but hardly any in regard to the proportion of statements with positive paranormal value. Hence one can distinguish between more and less 'risk-taking' psychics rather than between more and less 'good' psychics.

As regards the type of inductor in feedback conditions the following findings arise from the structural analysis:
(1) An object of the target person induces statements about the same types of topics as a photograph does. With both types of inductors psychics talk most often about personality traits (about 20%), psychological circumstances (about 15%), bodily health (about 13%), circumstances in work (about 9%), and about appearance and overt behavior (about 8%).
It further appears that psychics make some specific statements which are clearly inconsistent with each other when they talk twice about the same target person (without being aware of that condition), while they make hardly any specific statement which is consistent with another.
Moreover, despite a lot of feedback no psychic detected that the same person was the target person in series 1 and 2.
(2) Compared to a photograph an object induces more impressions about persons related to the target person. The psychic often considers an object as being a present or inheritance from a person related to the target person.
(3) A photograph induces more statements about favourable and unfavourable states of affairs than an object. An object induces more statements about neutral states of affairs.

As regards the amount of feedback, the following findings result from the structural analysis:
(1) The amount of feedback has no significant effect on the types of topics discussed in the statements.
(2) The number of topics about persons related to the target person increases when the amount of feedback decreases.
(3) Giving no feedback results in more statements about the past as well as about the future of the target person.
(4) Confined feedback induces relatively more statements about a favourable state of affairs compared to extended feedback as well as compared to no feedback.
(5) Giving no feedback results in many more statements preceded by a silence.
(6) Psychics are used to receive feedback in their daily practice. This appears from the fact that, in the no feedback condition they neglected a considerable number of times (14%) the information from the sitter that he was not acquainted with the target person, and instead made rhetorical rather than positive statements. It is important to note that all psychics had this tendency. The results also indicate that psychics have the tendency to elicit more, i.e.

extended feedback, in the condition in which the sitter gives confined feedback only. The important role of feedback in the behaviour of psychics is finally demonstrated by the fact that rhetorical statements significantly predominate in the first half of the session in all series except the series in which the sitter gives confined feedback only. In other words, in the case of confined feedback the psychics will try to elicit more feedback till the end of the session.

As regards the importance of the problem of the target person the following findings appear from the structural analysis:
(1) About 37% of the topics discussed in the statements of the sessions about the missing person were explicitly related to that problem. The distribution of the other topics discussed was about the same as those in the standard series, with the exception that in the case of the missing person temporal psychological circumstances (for example, difficult period in life) were discussed as often as general personality traits (for example, submissive), whereas in the standard series the number of statements about personality traits were heavily predominant. This suggests that a situation in which a person is missing is equally strongly associated both with temporal psychological circumstances of the person (for example, depressed mood) and with general personality traits (for example, anxious person).
(2) The number of topics about persons related to the target person decreases when the problem of the target person is of more importance.
(3) Contrary to what one might expect the number of explicit pieces of advice decreases when the problem is of more importance.
(4) The psychics are more inclined to consider more carefully what to say next when the problem is of more acute importance. This follows from the higher number of statements preceded by a silence.

As regards the set of statements having potential paranormal value compared to the set of statements with no potential paranormal value, most differences observed between the two sets of the standard series are present in the experimental series as well. The only remarkable difference is the following: in the set of statements with potential paranormal value, relatively more statements were preceded by a silence in the condition in which only an object was presented and in the condition in which the sitter provided no feedback at all than in the set of other statements. This difference was not found in the standard series. Thus, it appears that only in those (relatively difficult) circumstances the psychics use the silences more often to receive more spontaneous and specific impressions.

The statements with positive paranormal value and the statements with negative paranormal value are not different with respect to any of the characteristics studied in the structural analysis.

The results of the interactional analysis indicate that psychics most often react to a denial by giving a new interpretation to their impression and that the different experimental manipulations had only a slight impact on this type of behaviour of the psychics. The amount of information in the denial (extended plus confined feedback versus confined feedback only) does not influence the type of reaction by the psychic. Further, in the case of a missing person the psychics tend to react more often with accepting the denial and less often with giving a new interpretation.

ABSTRACT

The aim of the analyses presented in this paper was to study the effect of some variables on the content of sessions of psychics. These variables were (1) the type of inductor (photograph or object) in feedback conditions, (2) the amount of feedback by the sitter and (3) the importance of the problem of the target person.
It was found that the number of statements first of all depends on the amount of feedback by the sitter. Psychics offer more than twice as many statements when sitters give extended feedback than when they give no feedback at all. The variables studied hardly affect the individual preference of psychics for sessions of a certain length.

None of the variables had a significant influence on either the number of statements with potential value, i.e. spontaneous and specific statements, nor on the number of statements with positive paranomal value, i.e. spontaneous and specific statements which are correct. Moreover, only about 1% of all statements can be considered as possibly based on paranormal impressions. In view of the judging procedures applied, the low percentage is not supportive of a paranormal interpretation. This implies that it is very doubtful that research with psychics offers a more promising research method than the methods employed most often in parapsychology.
Psychics are rather consistent over the different experimental series as regards the proportion of statements with potential paranormal value in a session. From the size of variation in the proportion of statements with potential paranormal value among the psychics (3% to

15%) and from the size of variation in the proportion of statements with positive paranormal value (0% to 3%), it is concluded that it is more appropriate to distinguish between more and less 'risk taking' psychics than to distinguish between psychics in terms of more and less 'good' psychics.

The type of inductor (object or photograph) has only some minor influences on the structure of behaviour of psychics in feedback conditions. On the other hand the amount of feedback provided by the sitter has a considerable influence. In their normal practice psychics are used to receiving extended feedback from sitters. When the sitter restricts himself to confined feedback, the psychics try to elicit extended feedback by making a considerable number of rhetorical statements, i.e. statements in a question mode. This is even the case when the sitter informs them beforehand that he is unacquainted with the target person. When psychics are consulted for a serious problem, as in the case of a missing person, they are more inclined to consider more carefully what they say, to give fewer statements involving explicit advice and to react more often with the acceptance of a denial by giving another impression than with giving a new interpretation to a denied impression.

REFERENCES

Boerenkamp, H.G., Schouten, S.A. 'Estimating the potential paranormal value of verbal statements', J.o.P., 47, 1983, 121-130.

Boerenkamp, H.G. 'Potential paranormal value of statements of psychics acquired under feedback conditions', E.J.P., 5, 2, 1984, 101-124.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part I. Experimental design', E.J.P., 5, 4, 1985a, 327-348.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part II. The standard series', E.J.P., 5, 4, 1985b, 349-371.

EUR.J.PARAPSYCHOL., V.6, 1985

PARAPSYCHOLOGY AND THE EXPECTATION OF PROGRESS

John Beloff
University of Edinburgh

In an essay entitled 'Final Impressions of a Psychical Researcher' which first appeared in the American Magazine for 1909, the year before he died, William James made the following memorable pronouncement: (Murphy and Ballou 1961)
"Like all founders, Sidgwick hoped for a certain promptitude of result; and I heard him say, the year before his death (1899) that, if anyone had told him at the outset that, after twenty years he would be in the same identical state of doubt and balance that he started with, he would have deemed the prophecy incredible. It appeared impossible that that amount of handling evidence should bring so little finality of decision."

"My own experience had been similar to Sidgwick's. For twenty-five years I have been in touch with the literature of psychical research, and I have had acquaintance with numerous 'researchers'. I have spent a good many hours (though far fewer than I ought to have spent) in witnessing (or trying to witness) phenomena. Yet I am theoretically no 'further' than I was at the beginning; and I confess that at times I have been tempted to believe that the Creator has eternally intended this department of nature to remain baffling, to prompt our curiosities and hopes and suspicions all in equal measure,

..." and so on. However, James resists this temptation and continues as follows:

"It as hard to believe that the Creator has really put any big array of phenomena into the world merely to defy and mock out scientific tendencies; so my deeper belief is that we psychical researchers have been too precipitate with our hopes, and that we must expect to mark progress not by quarter centuries, but by half centuries or whole centuries."

It is now three quarters of a century since James penned these lines but James' lament is still echoed, again and again, by numerous parapschologists and others. Indeed it would be no exaggeration to say that nothing has so distressed the devotee or given so much impetus to the critic than the apparent lack of progress in this field. It is not that there has been any lack of evidence, some of it very impressive, it is that we have come to expect more from a science than the mere accumulation of facts and findings. With the advent of J.B. Rhine and his celebrated laboratory, hopes were kindled that this would inaugurate a new era of progress. The recent memorial volume was entitled: 'J.B. Rhine: On the Frontiers of Science' (Rao 1982) and the very first textbook of parapsychology which Rhine published with Pratt in 1957 was subtitled 'Frontier Science of the Mind' (Rhine and Pratt 1957). Rhine, indeed, always thought of himself as a frontiersman but the trouble is that a frontier that does not continually move forward is apt to degenerate into a mere fringe area. At all events, disillusionment eventually set in, parapsychology, it transpired, was, to use the apt phrase which Mauskopf and McVaugh chose as the title of their scholarly book on this period, 'the Elusive Science' (Mauskopf and McVaugh 1980). Today the problem of progress in parapsychology is still as troublesome for us as it was for William James. In this paper I shall discuss just what kind of progress is envisaged by those who lament the lack of it in parapsychology and whether it is reasonable or realistic to expect rapid or spectacular progress in a field such as ours.

Under the rubric of 'progress' then, a number of possible meanings need to be distinguished and I shall single out the three which seem to me to be most relevant. There is first progress in the understanding of the phenomena. This, I take it, is what James had in mind when he complained that 'theoretically' he was no further forward at the end of his career than he was at the beginning. Let us call this 'theoretical progress'. Next comes the question of credibility, recognition, status and so forth which may or may not depend on theoretical progress. The founding of the P.A. in 1957 was one such milestone, its affiliation to the A.A.A.S. in 1969 was another, the establishment of the Koestler Chair at Edinburgh University in 1985 is yet another and so on. I do not know what is the best name but let us call this 'sociological progress', it is progress in influence and power such as a political or religious movement would strive for but equally it is the objective of movements and schools within the history of science. Then, thirdly, progress may be assessed in terms of the practical benefits that accrue from a given science, we call this 'technological progress'. Hitherto, parapsychology has yielded little in the way of practical applictions. This may seem odd since, from the dawn of history, alleged paranormal powers were always used for some definite practical purpose. A parapsychological technology, however, would mean using the techniques developed in the laboratory to enhance psi functioning within some social context, be it commercial, medical, forensic, military or whatever. It is of interest to note, in this connection, that some parapsychologists, these days, notably those working on the West Coast, disheartened by our lack of success in trying to impress the scientific community with our experimental evidence, have turned, instead, to developing techniques for putting psi to work in just such real-life situations (Targ and Harary 1984). If their efforts prove successful then I think there is no question that this would make a much bigger impact on the public at large than any internal theoretical advances that we could expect. It remains to be seen, however, whether this recent development will have the outcome which its supporters are claiming for it.

Looked at from the broad historical standpoint, I think there can be no doubt that the idea of progress, or, at any rate, of progress as the touchstone of all human endeavour, is a relatively modern idea, a product of our Western civilization and of the scientific and industrial revolutions through which it passed. Francis Bacon stands out as one of the earliest prophets of the idea that knowledge means power but, more and more, the idea of progress came to be bound up in the public imagination with what we have here called technological progress. This reached its apogee in the late 19th century, an era of great optimism, when a spate of inventions and technical innovations was rapidly transforming the conditions of life. Since then technological progress has continued at an accelerating pace although it has created in its wake such vast social problems and has brought us so close to the total destruction of life on Earth that, while we have not ceased to believe in progress, we have shed something of the pristine Victorian optimism. Perhaps the most salient example of technological progress in recent times has been the way in which, for better or worse, computers have entered our lives at so many points. Actually technological progress, as such, is something that has been going on ever since prehistoric times, as every archaeological excavation can testify, but it was the peculiar and explosive combination of science and technology from the 19th century onwards that created the world in which we now have to live.

When we consider theoretical progress, however, the picture becomes somewhat more complicated. The old idea of science progressing in a piecemeal, linear fashion, with each new generation of scientists adding their quota to the achievements of their predecessors, is one that has taken some severe knocks since the downfall of classical physics. The newer post-Kuhnian graph of scientific history resembles more a zig-zag than a staight line vector or an exponential curve. Arthur Koestler expressed this very vividly when he said: 'The progress of science is strewn like an ancient desert trail with the bleached skeletons of discarded theories which once seemed to possess eternal life.' (Koestler 1976). But, however ruthless science may be in what it rejects, as one paradigm succeeds another, there always seems to be a net gain in overall sophistication. Another way of putting this would be to say that scientific revolutions move in spirals rather than circles, there is always some movement in the upward direction however deviously achieved.

Ever since Thomas Kuhn published his The Structure of Scientific Revolutions in 1962, parapsychologists have pondered the idea that what was delaying progress in parapsychology was that it had got stuck at the 'preparadigmatic' stage. Sooner or later, it was suggested, some scientists of genius, some new Einstein, would come along whose startling new insights would show us how our present confusion could be resolved and how parapsychology could at last graduate to becoming a normal science (McConnell 1983; Hovelmann 1984). How plausible is this idea? The point I want to make, and it is the main point I want to make in this paper, is that such an idea, though superficially so appealing, makes sense only on the assumption that parapsychology is essentially a physical science. Some attempts were made to extend the Kuhnian schema to the social, behavioural and psychological sciences but they did not work. It transpired that none of these sciences were paradigmatic in the Kuhnian sense. In economics, anthropology, sociology, linguistics, to say nothing of psychology, there has never been a single universal paradigm; at most certain schools of thought or certain theoretical models were dominant at certain times (Mackenzie 1977). Accordingly, progress in these fields has always meant something very different from progress in the physical sciences. They may have given birth to new ideas each with a certain fertility and lease of life, to new methodologies that take advantage of new mathematical techniques or of new instrumentation, but we would be deceiving ourselves if we were to expect that these sciences could revolutionize our understanding of human behaviour or human nature in the way that the physical sciences have revolutionized our understanding of the physical universe and of the behaviour of matter. For one thing we already know too much about human beings.

There are, at the present time, as you are well aware, a formidable body of thinkers within the parapsychological community who have come to regard parapsychology as an extension of quantum physics and psi phenomena as special cases of physical phenomena. If they are correct then the progress of parapsychology is assured since its future is then bound up with the future of physics from which it can derive support and to which it can contribute. The question is, of course, are they correct? No one knows and, perhaps, only time can tell. It is difficult for a lay person like myself to judge; presumably the final arbiters must be the physicists themelves. I have elsewhere expressed my misgivings on various logical grounds (Beloff 1985) but I am not here concerned to argue the case but merely to point out the connection between the physicalist view of psi and the expectation of progress. For, of I am correct, if, that is, psi phenomena are essentially mental phenomena and parapsychology is essentially a psychological science, then our expectations of progress must perforce be far more modest. For it could well turn out to be the case that the very spontaneity and elusiveness of the phenomena that have for so long resisted the scientific approach, with its emphasis on predictability and controlability, may be an inherent feature of these phenomena in virtue of their being manifestations of mind. I would no more expect psi phenomena to conform to certain universal and immutable laws than I would expect human intentions and volitions to be explicable in terms of the electro-chemistry of brain cells.

One so often hears it said both inside and outside the parapsychological community, that unless we can develop some theory or theoretical framework so that the phenomena we study would no longer appear as isolated anomalies, we shall never gain credibility no matter how long we go on piling data upon data. This may be true, just as it may be true that parapsychology is destined to remain forever the esoteric interest of a minority. I would like to think, however, that the time will come when it forms part of the outlook of all educated people in our society. The question that concerns me, therefore, is what strategy is best calculated to bring this about? And here I feel that the pursuit of theoretical progress may constitute a diversion from the more immediately realizable goal that I have called sociological progress. By all means let there be theoretical speculations but the belief that this will solve the problem of the paranormal or normalize psi phenomena to the general satisfaction of the scientific community strikes me as wishful thinking. At the same time, I do not think that the more direct method of accumulating data has yet been exhausted. In particular, I feel that the latest research reports from the laboratory at Princeton University under Dr Robert Jahn offer new promise. What I find encouraging in their approach is the realization that, to make any impact at all, what we need is large quantities of data. To get a highly significant result in a one-off experiment with some subject who is never again available is never going to be very convincing. Accordingly, they have made it their policy to accept anyone who is willing to spend sufficient time in their laboratory and agrees to perform according to their stated protocol. In terms of the sheer bulk of their data, accumulated over a period of some five or six years, and in terms of the stringency of their methods and safeguards, they have succeeded in setting a new standard of excellence and, in the event, have been rewarded with highly significant results, both in their ESP research on Remote Perception and in their PK research with REGs. (Dunne, Jahn and Nelson 1983; Nelson, Dunne and Jahn 1984) The only reservation that one feels called upon to make at this stage is that there is so little else in the field that comes up to such standards and, in a field where experimenter effect has proved to be so ubiquitous, one would like to see similar results from other laboratories. Jahn and his associates are not, of course, oblivious to the desirability of theories, in a rather vague sort of way they suggest that at the present time quantum theory may well provide the best clues in seeking an explanation of their findings (Jahn and Dunne, 1983), but it is their data, not their theorizing, that give us the ammunition that we need so that, with more such data, we should be ready to go over to the offensive and challenge our critics either to accept the evidence or produce a reasonable counter-explanation.

Proponents of the theoretical progress like to assume that facts can be of no interst in themselves but only insofar as they advance our theoretical understanding. This assumption, however, stems from a too exclusive preoccupation with the natural sciences. It could be argued that the facts of nature concern us only when they are either useful for us to know or when they illustrate some theory which has intrinsic intellectual appeal. It is quite otherwise, however, in what we may loosely call the human sciences. Thus, historical or biographical facts, for example, may be devoid of any utility and, equally, may fail to exemplify any general law or principle. Nevertheless, if they are dramatic enough, they engage our attention just because, as human beings, we are naturally curious about the human condition. If, now, parapsychology is to be included in the human sciences anything it can tell us about the incidence or extent of psi abilities or psi events cannot fail to be of interest irrespective of the particular theoretical interpretation we may prefer. For example, the compiling of a scholarly catalogue of reincarnation type cases on a global scale represents one important kind of progress (Stevenson 1975, 1977, 1980, 1983). But, in the last resort, it would be hard to imagine any discovery in parapsychology that could equal in importance the unequivocal demonstration that psi exists, that is, the bare fact that human beings can, on occasion, communicate in a manner that transcends the constraints of space, time and matter. And I say this precisely because so much in contemporary science and philosophy is predicated on the denial of this fact. What revolution in science could have a greater effect on Western culture than its taking of parapsychology on board as an authentic body of knowledge? For, whichever way one looks at it, a world that includes psi events demands a very different world-view from one that excludes them. Perhaps, therefore, the most important measure of progress we can adopt is in terms of the success attained in the pursuit of that goal.

REFERENCES

Beloff,J. 'Parapsychology and Radical Dualism', Journal of Religion and Psychical Research, 8, 1985, 3-10, Abstract in Research in Parapsychology 1983, 1984, 39-42.

Dunne, B., Jahn, R., and Nelson, R. 'Precognitive Remote Perception', Technical Note PEAR 83003, 1983.

Hovelmann, G.H. 'Against Historicism: Critical remarks on Thomas Kuhn's conception of science and its reception in parapsychology', J.o.P, 48, 1984, 101-119.

Jahn, R.G. and Dunne, B.J. 'On the quantum mechanics of consciousness, with application to anomalous phenomena', PEAR 83005.1, 1983.

Koestler, A. 'On truth and beauty (1976)', in Koestler: Bricks to Babel, London, Hutchinson, 1980, p. 369.

Mackenzie, B.D. 'Behaviourism and the limits of scientific method', London, Routledge and Kegan Paul, 1977.

McConnell, R.A. 'An introduction to parapsychology in the context of science', University of Pittsburg, McConnell, 1983, see Chap 18.

Mauskopf,S.H. and McVaugh, M.R. 'The elusive science: Origins of experimental psychical research', Baltimore, MD, Johns Hopkins University Press, 1980.

Murphy, G. and Ballou, R.O. (eds.) 'William James and psychical research', London, Chatto and Windus, 1961, see p. 310, (also published by: New York, Viking Press).

Nelson, R., Dunne, B. and Jahn, R. 'An REG experiment with large data base capability, III: Operator related anomalies', Technical Note, PEAR 84003, 1984.

Rao, K.R. (ed.) 'J.B. Rhine: On the frontiers of science', Jefferson, N.C., McFarland, 1982.

Stevenson, I. 'Cases of the reincarnation type', Vols. I-IV, Charlottesville, VA, University of Virginia Press, 1975, 1977, 1980 and 1983.

Targ, R. and Harary, K. 'The mind race: Understanding and using psychic abilities', New York, Villard Books, 1984, see Part III 'Psi in everyday life'.

J. Beloff
Department of Psychology
University of Edinburgh
7 George Square
Edinburgh
Scotland EH8 9JZ

EUR.J.PARAPSYCHOL., V.6, 1985

THE DUALIST TRADITION OF PARAPSYCHOLOGY

Hoyt Edge
University of Utrecht

Sometimes more progress can be made in a discipline by stepping back, by assessing where one is by seeing one's origins, than it is by suggesting new theories. It is for this reason, essentially, that I write this small piece. A second reason is more pedagogical -- often it is difficult among the welter of professional writing to find a position stated clearly and without frills to allow the student, and perhaps even the professional immersed in the detail of daily activity, a context in which to understand the contemporary position of the field. Both of the these justifications merge in the hope that clarity about simple and basic matters will be useful. While aiming at such simplicity I will certainly not provide much that is new to those practitioners who have spent time assessing the tradition of parapsychology, but perhaps others, because unfamiliar with the histories of parapsychology, philosophy, and science, can find something useful in this essay and gain a perspective on the field. Therefore, I will provide no new answers to important questions; my job is rather an analysis of the world view behind traditional parapsychological research.

I

Most of the early workers in the field of psychical research were people committed to a dualistic ontology -- the view that the world is composed of two radically different kinds of realities, mind and matter -- and at the same time they were also dedicated to the scientific method. Others have written about the commitments of the founders of the Society for Psychical Research (see, for instance, Gauld, 1968), so I would like to concentrate on more recent investigators, particularly on J. B. Rhine, who is an especially good example because he not only shaped the field since 1935, but he was very interested in wider philosophical questions. Indeed, from Louisa Rhine's bibliography of her husband (and inevitably of herself, also), as well as Mauskopf and McVaugh's history of the early days of parapsychology, we are led to see the burning religious questions which motivated the intelligent and inquiring mind of the founder of modern experimental parapsychology. For a while he became convinced of materialism, holding it "as kind of worship, comparable to anything I had experienced in my period of response to the Christian religion" (quoted in Mauskopf and McVaugh, 1980, p. 72). However, reading Bergson's 'Creative Evolution' converted him to a position which he held the rest of his life -- a belief in dualism in which the mental can be investigated scientifically through research in parapsychology (note 1).

It is not difficult to garner quotes to show that it was Rhine's view that parapsychology is opposed to materialism. In 1966, he defined a parapsychological phenomenon as one which "is beyond physical principles of explanation" (p. 84). Talking about PK in 'Reach of the mind', he says that it is "nonphysical" (Rhine, 1947, p. 116). Later in this book, he says that "psi researches show that the natural human mind can escape physical boundaries under certain conditions;" (p. 205) and in asking: "Is there anything extraphysical or spiritual in human personality?" he responds: "The experimental answer is yes" (p. 206). For Rhine, parapsychology clearly showed that "man is something more than a physical being," (p. 209) and that the "nonphysical activity of the mind is demonstrated" (p. 214).

Of course, Rhine is not the only one to hold this position in our edition. Whately Carington (1933) wrote a book called 'The Death of Materialism' in which he garnered all sorts of scientific and philosophical argumentation against physicalism. In contemporary parapsychology, several researchers have employed the dualistic position of the neurophysiologist Sir John Eccles in support of their dualist position. Although he does not use Eccles this way, nevertheless Ed Kelly (1979) describes Eccles' position nicely: "His dualism is apparently founded chiefly upon the Sherringtonian intuition that the phenomenon of the brain and mind are irreducibly incommensurate, but it also appears to have sources in his catholicism, in a powerful experience (an OBE?) he had as a teenager, and in a naive belief that only dualism can restore to the human person the sense of wonder, of mystery and value" (p. 374). This quote focuses our attention on several issues which traditionally have been interrelated: psi phenomena, dualism and value. What is the intellectual background which gave rise to the context of thinking in which the marriage of these ideas seems so perfect?

II

As has been pointed out by others, that context is the Cartesian view of the world. As a scientist Descartes was impressed with what appeared to be the universality and certainty produced by science, at least in his own field of mathematics, and he wondered why philosophy, following the methods of science, could not come out from under the domination of authoritarian religion and achieve the same universality and certainty about important philosophical questions: the nature of the self, the existence of God, the foundation of morality, the hope of survival, and the meaning of human life. The aim, therefore, was to be as scientific about these questions as possible, although Descartes believed that they could not be directly tested empirically, as Galileo was doing with the movement of heavenly bodies.

There was a revival in the sixteenth and seventeenth centuries of the atomism of the ancient Greek thinker, Democritus; thus, in rejecting the medieval world, scientists of all sorts were formulating theories in atomistic and mechanistic terms, as Galileo was doing in support of Copernicus. In this context, Descartes' important contribution was that he was one of the first to apply mechanistic theory not simply to large-scale heavenly bodies but to the individual human body (which, incidently, caused his work to be banned by the Church after his death). Descartes had seen and been impressed by the French Royal Gardens which had been planned around a number of statues that were either themselves movable or had parts which were movable, driven by the flow of water set off by the weight of a person strolling along the walk. Evidently not only did these statues have some sort of sound system attached to them and so appeared to talk, but they moved in seemingly purposeful ways, acting 'embarrassed' or 'angry' in appropriate contexts. So impressed was Descartes by this display of mechanical movement (and undoubtedly for other reasons) that he asserted that the body was nothing but an automaton, a mechanical device drivable under its own power. Of course, for the human there had to be something non-material, something which could think, which prescribed the movements of this non-thinking automaton, and this was the individual mind. I will not discuss the details of his argument for the existence and nature of mind and matter (the body being an exemplification, or mode, of matter), but suffice it to say that he arrived at the conclusion that the world (and specifically each person) was composed of two radically different sorts of things -- two sorts of substances, to use his language. A substance was something which was wholly independent and which differed in essential ways from any other substance. Let me describe these two Cartesian substances in a schematic way:

MIND	MATTER
Thinking	Non-thinking
Non-spatial	Spatial
Seat of value	A-valuable
Free	Determined
Purposeful	Mechanical
Private	Public
Subjective	Objective

This division of the world into two non-overlapping regions served another function. Galileo had gotten into trouble with the Church because of his scientific theories, and the scientific part of Descartes believed that science should have free reign in its endeavors. Dividing the world into two regions facilitated this wish, with science having authority in the area of material reality, and the Church having authority in mental or spiritual matters. But this division not only gave science its freedom, it also bound it to a certain conception of matter and to a certain methodology adequate to investigate that kind of matter, i.e. an atomistic and mechanical investigation of non-thinking, a-valuable atoms in space which followed universal deterministic laws. And science has been defined essentially in this way since the seventeenth century.

To say the obvious, science has followed this approach not only because of the Cartesian philosophical stance but because, using this method, it has been so successful in mastering the world. Saying this, however, does not imply that another approach might have been more successful; nor does it ignore the obvious fact that we find what we look for. If we look for a mechanistic, atomistic world, there is no reason to think that we will find anything else, except by accident. But it must be acknowledged that within this traditional scientific viewpoint, we can point to remarkable progress and accomplishment.

Since the mind was not a direct object of scientific investigation -- in fact, it was precisely the kind of entity which science did not explore -- the attempt to inquire scientifically into the function of the mind was much later in coming. But one can already see the tension involved in such an inquiry: How is science, dedicated to investigating public objects, supposed to inquire into non-material mind? Thus, questions about the mind were traditionally reserved for philosophy and religion. When science finally focused on the mind, however, its methodology was already clear; if science was to investigate the mind using its accepted methods, it must assume that the mind is determined and mechanical. Two general approaches can be taken: one is to ignore the mind and assert that the mental can be explained scientifically in terms of the physical (identity theory, epiphenomenalism, behaviorism), and the other is to assert that the mental is not in space and therefore not material, but discuss it otherwise as if it were physical. Thus, as an example of the second approach, Freud described the mind in the mechanistic language of instincts, and the pull of drives and also admitted that what he was offering was a naive theory to be replaced later by a more sophisticated and complete neurophysiology. As an example of the first approach, Skinner went further, accepting the description of both mind and matter outlined by Descartes, and arrived at behaviorism. If one wants to investigate the person scientifically, one can do so only using science (that seems pretty obvious), but tradition dictates that science can investigate only that which is spatial, mechanical, deterministic and a-valuable. Hence, on one level Skinner seems to be saying that there is no mind, but in his more careful moments, he simply says that the mind is irrelevant for science. If it exits, it is not the kind of thing which could be investigated by science. Therefore, it matters little whether Skinner is a materialist or an epiphenomenalist (i.e. one who says that mind exists but it is an impotent by-product of physical processes); both imply that mind is irrelevant for science.

Although it goes beyond the purpose of this paper, it may be mentioned that cognitive psychology, the newest approach to the problem, does not fare much better. While it was developed specifically as a reaction to behaviorism, urging the inappropriateness of ignoring the mind, nevertheless, it lies also in the same tradition. As opposed to behaviorism which says that we cannot investigate into the black box which occurs between stimulus and response, cognitive psychology argues that we can, deducing indirectly what must be happening within the black box under certain circumstances (by techniques such as measuring reaction time and postulating that one cognitive strategy must have occurred rather than another since it would have e.g. taken more time than the other one). Although we cannot directly study what is not public, we can do so indirectly; yet the mind is assumed to be mechanistic and rule-governed, and the primary metaphor used is of the mind as a program of a computer (note 2).

III

That parapsychology historically grew out of this Cartesian context should be obvious. On the one hand, psi phenomena have been understood in terms of this dualism -- indeed, we have seen that Rhine clearly defined psi in these terms (note 3). Telepathy, for example, is conceived to be communication from one mind to another (see Figure 1), a process entirely non-physical. Somehow communication goes directly from one mind to the other without a physical intermediary or instrument (necessarily so, Rhine thought, since there was no sensory organ to receive it).

That Rhine conceived of the mind in the same terms described by Descartes can be seen further from a series of quotes, all taken from The Reach of the Mind (1947).

1) The mind is different from the body -- "By the discovery of an experimental sanction for the psychocentric conception of man we can be brought to think of people all over the world as being more than bodies" (p. 222).

2) The mind is non-spatial -- "When ESP was found to function without limitation from time and space, this discovery was taken to mean that the mind is capable of action independent to some degree of

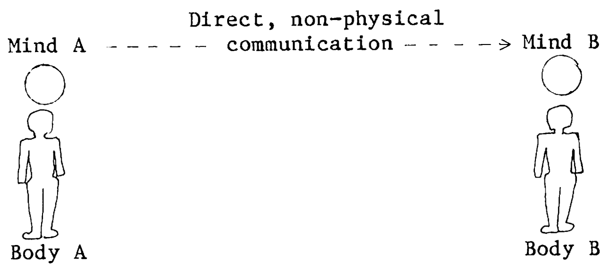

Figure 1

the space-time system of nature" (p. 213).

3) The mind is purposeful and not mechanical -- "PK results do not follow the laws of mechanics," (p. 116) and "PK is plainly a purposively oriented operation and therefore cannot be physical" (p. 117).

4) The mind is free -- "The characteristics of intelligent, volitional action in human behavior are so different as to be in some respects contrary to those of the causal principles basic to the physical sciences." (p. 117) Further, "For only if the will has a nature different from the world on which it acts can it operate on free principles--free, that is, from the laws of the other systems. Certainly, if we may say anything about precognition, it is that it does not conform to physical laws. If, then, the mind can predict events, it is just that much more free from, because it is different from, the physical system upon which it acts" (p. 84).

5) The mind is the seat of value. This proposition has always been tied to the notion of freedom, as philosophers have argued that it is difficult to conceive of there being ethical behavior (performing the right action and following the good) without freedom of choice and action. We see these two ideas connected in Rhine, who even referred at one time to the "problem of ethical freedom" (p. 220). Further, he states: "The more we are led on the one hand to think of our fellowmen as deterministic, physical systems--robots, machines, brains--the more heartlessly and selfishly we can allow ourselves to deal with them" (p. 219). A few pages later, he returns to the same theme, saying "the acceptance by man of a free ruling

factor over and above his physiochemical system has raised him from the brutish, selfish, quarrelsome impulses of his primitive nature" (p. 223).

I believe that these quotes justify our saying that Rhine, and thus the discipline following him, have traditionally conceived of both the world and parapsychology in Cartesian terms.

In addition, Rhine believed that the scientific method was universal, and so all important questions (even religious) had to be approached scientifically. But given the Cartesian context he accepted, a paradoxical situation arose -- because of both the methodology it was committed to as well as the nature of the objects it investigated, it looks as if science was given the job of proving realities which contradicted its methodological and ontological commitments. Seen from this context, in contradiction to 'normal' psychology which redefined mind to investigate it as if it were a determined, mechanistic phenomenon, parapsychology completely retained the Cartesian definition of mind. Therefore, parapsychology becomes the scientific investigation into mind which Descartes had excluded from science and placed in the realm of religion and philosophy. In these terms the job of parapsychology is to investigate non-spatial, free, purposeful, value-laden phenomena using a method devised to investigate spatial, determined, mechanistic and a-valuable objects. Since the phenomena don't match the method, assuming the traditional context, it is no wonder that traditional science has looked askance at the scientific status of parapsychology, as it is difficult to see how one can understand psi phenomena using a methodology committed to describing exactly the opposite kind of object.

One can see now that the thrust of the a priori arguments against parapsychology (and most of the others implicitly) also assume this traditional Cartesian context. From this perspective, science is simply not in the business of taking traditionally understood mental phenomena into account; its commitments are to other phenomena. The criticisms of the Humanists can be understood in this light, as they, too, accept the traditional context. To them, since science is committed to describing the world in non-mental terms and has been so successful in doing so, psi phenomena should not be taken seriously. Within the traditional context, mental phenomena are not objects which science could prove, thus parapsychology is engaged in a false start.

This tension between the objects of parapsychology and its method can be pointed to in another way by considering a quote from Gregory Bateson. He says (1979) that ESP, materializations, OBEs and spiritualism in general are "all symptoms, mistaken attempts at cute efforts to escape from a crude materialism that becomes intolerable. A miracle is a materialist's idea of how to escape from his materialism." (p. 210) What he is saying is that if one remains within the traditional context in which the world is described by science as (crudely) material, the only way out is to assert a miracle -- one can only assert that something beyond the system, which is not scientifically understandable, impinges upon the system and affects it in a way which cannot be understood by normal means. Science (the material world) and parapsychology (the mental world) are exclusive in the conventional context; there really is no adequate way to bring together two so disparate realms.

IV

Given the traditional context, it looks as if the most parapsychology can do is to try to prove that an anomaly exists within the system. That this seems to be the primary thrust of the discipline -- in spite of all the process oriented experimentation -- can be seen by noting that the definition of psi is purely negative (i.e. those phenomena which don't seem to fit into the traditional context). What this means is that we are defining psi in terms of a tradition to which we have an ambivalent relationship. On the one hand, we view psi phenomena as being mental, as Descartes defined them, but on the other hand, we reject the view that these phenomena must remain in the domain of religion and philosophy; thus we attempt to employ a science, developed to deal with deterministic, mechanistic material phenomena, to investigate a radically different sort of reality. But strange consequences seem to follow from this ambivalence. For instance, in a PK experiment, we must do our best to set up the experiment so that there is no physical influence present, so that no physical energy can go from point A (subject) to point B (target). Let us assume that we are successful in getting a result which is interpreted as PK. However, since we have excluded physical influence, how are to to interpret what has happened? The traditional context argues that all influence is caused by some energy. Since we have excluded physical energy, parapsychologists proceed to talk about mental energy or influence, because we employ the methodology, terminology and modus operandi of the traditional context which is aimed toward understanding matter. Thus we try to combine the two realms by assuming a hybrid -- mental energy or influence. But this brings its own problems. Normally, we want to say that any theory (or idea within a theory) must be falsifiable if is to be considered scientific. Therefore, if one would want to ask, for instance, what the difference is between an influence accomplished though physical energy and one accomplished without such energy, the traditional way to answer the question is to set up an experiment which excludes the possibility of physical energy. But mental energy has a peculiar property -- it cannot be excluded by any known means (there are no known limits to psi); it does no seem to be falsifiable. It does not seem that we can ask: Was this instance of PK a result of mental influence or not? There does not seem to be a way of experimentally differentiating the two concepts -- all parapsychology can do is exclude the physical influence. Why not simply say that there is no influence in PK -- simply a correspondence -- rather than saying that there is a mental influence? If PK is a non-instrumental effect -- that is, if no instrument can be postulated or studied -- all that is left is a correspondence.

To assert merely a correspondence, however, would take us out of the traditional context in which all action occurs by means of influence. The result seems to be that since science is dedicated to investigating and explaining one kind of phenomenon and parapsychology is dedicated to proving anomalies in that system, it becomes impossible to explain psi phenomena scientifically. How can one explain non-spatial (etc.) phenomena using a system of explanation aimed at spatial (etc.) phenomena?

Ironically, the way out of this difficulty was also implicit in The Reach of the Mind, the same book we quoted from so extensively above in proving Rhine's acceptance of Cartesian ontology. He said (1947), speaking of ESP: "It does not depend upon any kind of physical principle known, or likely to be derived from the space-time physics of the present day" (p. 204, italics mine). In other words, one opts out of the traditional context. This is, of course, why so many contemporary parapsychologists have offered new theories, such as the field theory, the conformance behavior model and the observational theories (or at least in some interpretations of the latter) in which several of the traditional characteristics of matter have been qualified in fundamental ways.

What may not be so obvious, however, is the implication of such a move in parapsychology (and certain areas of normal science). To put it boldly, the evil of Materialism which parapsychology set out to undercut depends on a commitment to a certain definition of matter (and science), and if that definition (and science) is changed, it is not obvious that the same questions need to be asked, philosophically or scientifically. For instance, if the new context for investigation is one in which the basic material (may I still use this word?) of the world is no longer exclusively described deterministically, mechanistically or atomistically, do we need to be concerned with proving the existence of a mind which is purposeful and free? In other words, in our tradition a certain view of mind has generally gone hand in hand with a certain view of matter, and if one concept is changed, the other is correspondingly called into question. It is no longer obvious, and I think many if not most parapsychologists see this already, that parapsychology has anything to say about proving the existence of mind, freedom or value. We cannot automatically assume that "the age-old problem of freedom comes under the light of research" (Rhine, 1947, p. 220) until we are clear about how we describe the world. There may be no problem of freedom we are faced with! To give an anthropological example, the Balinese are not bothered by any 'problem of freedom.' They do not view matter as being composed of inert atoms, lifeless, a-valuable and determined -- something to gain and exert control over -- but nature is enchanted with spirits (some of which are not too bright: a useful concept, I think). Nor do they sense a lack of freedom in their social organization. In spite of a rigid and highly prescribed system of expected behavior, they do not feel constrained by it but rather find security in knowing how others are going to act, the greater fear being of the unknown than of constraint. Thus, neither metaphysically nor socially is freedom a problem; hence, they not only do not talk about freedom but they do not conceive of the world in such a way that the concept is totally applicable.

V

In the beginning of the paper, I indicated that my aims were limited. I do not wish to speculate on a new context for parapsychology; there is already much discussion in the literature doing that. My purpose was more looking back than forward, but a few lessons can be learned from this exercise. Perhaps the most important one is that a reconceived parapsychology may form a break with tradition, not only conceptually in rejecting Cartesian ontology, but in the motivation for practicing parapsychology. Depending on the new context, parapsychology may have much or little to say about what it means to be a person. It may have nothing at all to say about religious and ethical questions (unless these in turn are radically reconceived). Whereas once it may have been adequate to view the task of parapsychology as forming a strong bulwark against the evils of materialism, its task may be more managable now. Whereas parapsychology once wanted to undercut science, perhaps now a reconceived version can be more secure in its scientific aspirations.

NOTES

1. The fact that he later came to espouse a position which he called 'relative dualism' does not in any way undercut my analysis of the conventional context of parapsychology. Since a closer examination of Rhine's interesting position would not further the clarification of the parapsychological tradition I am attempting here, I will not pursue it.
2. At the other extreme, humanistic and transpersonal psychologies assert the reality of the mind but at the expense of being rigorously scientific, as these areas are related closer to clinical than experimental psychology. Phenomenological psychology tries to cut the Gordian knot by denying the dualistic tradition.
3. Although there is some indication that he did not feel entirely comfortable with the Cartesian version of interactionism (but even Descartes recognized the unsatisfactory nature of this position), Rhine was given the opportunity to reconceptualize the context of the radical distinction between mind and matter when he took a course from Whitehead at Harvard, but Rhine stopped attending the class, saying that Whitehead's non-dualistic position was "far too abstract to be useful" (quoted in Mauskopf and McVaugh, 1980, p. 78).

REFERENCES

Bateson, G. ' Mind and nature', New York, E. P. Dutton, 1979.

Carington, W. 'The death of materialism', London, George Allen and Unwin, 1933.

Gauld, A. 'The founders of psychical research', London, Kegan and Paul, 1968.

Kelly, E. (1979) 'Converging lines of evidence on mind/brain relations', In: B. Shapin and L. Coly (Eds.), 'Brain/mind in parapsychology', New York, Parapsychology Foundation, 1979.

Mauskopf, S. and McVaugh, M. 'The elusive science', Baltimore and London, The Johns Hopkins University Press, 1980.

Rhine, J. 'The reach of the mind', New York, William Sloan Associates, 1947.

Rhine. J. 'ESP-what can we make of it?', J.o.P., 30, 84, 1966.

Hoyt L. Edge
present addres: Rollins College
Winter Park, Florida 32789
U.S.A.

EUR.J.PARAPSYCHOL., V.6, 1985

A POSSIBLE ARTIFACT IN A PK TEST FOR BABIES

Tomasz Troscianko
Susan Blackmore
Brain and Perception Laboratory
University of Bristol

In a recent issue of the 'European Journal of Parapsychology' Dick Bierman (1985) reports two series of computer controlled experiments in which targets were generated by a Random Number Generator (RNG) and the subjects were babies. The main prediction was "There will be a differential effect between feedback and non-feedback trials" (Bierman, 1985, p.377).

In the first (Amsterdam) series, with two babies, overall results were close to chance and this prediction was not confirmed. A second series was run in Bristol, using the same software and using one baby as subject. The present authors were the experimenters. This confirmed the main hypothesis showing a significant feedback effect with significant hitting in the feedback condition.

In an experiment of this kind the conclusion that psi was responsible for the results depends crucially on the randomness of the RNG, especially since there was no control condition without a baby present. We were concerned about possible non-randomness and therefore set about testing the RNG using the actual computer and RNG electronics used in the original experiments.

Long series of randomness tests showed no overall first order biases. However there appeared to be possible irregularities when the computer was 'cold' and we decided to investigate these further.

We thought it possible that temporary non-randomness may occur which is dependent upon the state of the computer at any given time. For example the output of the RNG might be influenced by the thermal state of the computer and by input/output activity occurring immediately prior to number generation. All these might be expected to vary the operating characteristics of the Zener diode used in the RNG circuit. The thermal state may affect breakdown voltage, input/output activity the supply voltage. In addition, there might be important interactions between these two parameters.

There are obviously differences in input/output activity between the feedback condition (in which a tune is played and a picture displayed) and the non-feedback condition in which there is no sound or display. If such effects were temperature dependent then they might be expected to be greater if the computer were run 'cold', that is having only recently turned on. It would then be in its least stable state. Many of the trials in the original series were run with the computer 'cold'.

To test this hypothesis we ran two more series of trials using the same system and procedure as before only with no baby present. To rule out any psi-mediated feedback effects by others all trials were run out of sight and ear-shot of anybody else. There were 30 trials run with the computer 'cold'; that is, it had just been turned on and had previously been off (but at normal room temperature) for at least 12 hours. There were 30 trials run with the computer 'hot'; that is, it had been on for more than 6 hours prior to the test.

We predicted that the positive results of the original experiment would be replicated in the 'cold' computer condition, but not necessarily in the 'hot' condition.

Biermans's software produces two scores for each trial with MCE=8 (feedback and non-feedback). For the 'hot' computer mean scores were 7.23 and 8.13. These are not significantly different from MCE and there is no significant difference between them.

For the 'cold' computer mean scores were 8.13 and 6.7. There is no significant hitting in the feedback condition but there is the expected differential effect, with the feedback score significantly higher than the non-feedback score (t=2.21, df=29, $p<.05$ two-tailed). Our prediction is confirmed: a 'Bierman effect' for the 'cold' computer but not for the 'hot' computer.

We infer that some kind of temperature-dependent artifact may be responsible for this differential effect. Of course it is possible to argue that some kind of psi-mediated experimenter effect of expectation was responsible, but this seems purely ad hoc and unhelpful since the state of the computer (hot or cold) appears to influence the outcome and this was not embodied in Bierman's original predictions nor any rational extension of them.

We make no specific suggestions as to the mechanism of this possible artifact beyond the original speculations which prompted us to run the experiment. We only wish to point out that these results suggest that in PK experiments with RNGs non-randomness (i.e. malfunction of the RNG) may appear in the actual running of a program while being undetectable in usual randomness tests run over a long period (e.g. all night). Also, if we have detected non-randomness in the behaviour of the RNG, it may well only be detectable in a short series, as in a real experiment, not in millions of randomness trials.

We conclude that tests for non-randomness may need to be specific to the situation (e.g. program and computer) in which the RNG is being used. Also if claims are to be made about the role of the subjects in PK experiments, control trials without subjects are essential.

REFERENCES

Bierman, D.J. 'A retro and direct PK test for babies with the manipulation of feedback: A first trial of independent replication using software exchange", E.J.P., 5, 1985, 373-390.

Tomasz Troscianko
Susan Blackmore
Universitaets-Augenklinik,
Schleichstrasse 12,
D-7400 Tubingen,
West Germany

EUR.J.PARAPSYCHOL., V.6, 1985

AN IMPOSSIBLE ARTIFACT

Dick J. Bierman

1. Introduction

The arguments put foreward by Troscianko and Blackmore (see this issue pages 95-98) to 'explain' the Baby-PK testdata as an artifact are weak. The effect found in their cold versus hot randomness tests does not account for the main effect (which they label the Bierman-effect) of the original experiment. Moreover the suggested physical (mal)function of the apparatus cannot (even theoretically) cause the effect which was found in the experiment.

1.1 The experimental and the control data

The experimental data consists of the Bristol-series, 33 runs of attempted PK with a baby present at the RNG-system. Troscianko and Blackmore report on 30 runs in a 'cold' and 30 runs in a 'hot' condition while there was no subject present. However inspection of the LOGFILE (see E.J.P., 5, 4, p.380) shows that in fact more control runs were performed. It concerns 3 runs in the 'cold' and 8 runs in the 'hot' condition. For the omission of the latter runs is some justification since they were run with very short intervals. However for the omission of the 'cold' runs there is no explanation. Apart from this slight carelesness, it is highly questionable of such a small number of runs constitutes a serious control and there is no justification given for the actual number of 30 which seems rather arbitrary. (See for a more general discussion of this topic paragraph 6)

2. The prediction for the control runs

The prediction that there would be an artifact in the 'cold' condition appears ad or post-hoc since for the specific brand of computers used in this study, and for electronic circuits in general, problems sooner are to be expected at high temperatures. There are even special fans made for this type of computer to prevent them from high temperatures. Prediction of malfunction at high temperatures would have been more in line with this commonly known phenomenon.

3. The data

The data on which the analyses for the 'Bierman' effect are reported are incomplete and the analyses are in error. Not only are some runs omitted from the data but also the data for two (of the four) conditions which are an intrinsic part of the 'Bierman' effect were omitted. The authors give a rationale for this omission (see paragraph 4) but since their conclusion is . . "our prediction is confirmed: a 'Bierman effect' for the 'cold' computer" . . , they should use all conditions, since the 'Bierman' effect concerns all 4 conditions. After correction for this selection the results for the 30 control runs become:
Number of hits in FB = 306 versus number of hits in NFB = 272 (chi-square=2.01, n.s.)
This should be compared with the experimental data of
Number of hits in FB = 377 versus number of hits in NFB = 314 (chi-square=5.586, $p<.02$).
If one accepts their rationale to select only two conditions then their data are:
Number of hits in FB = 244 versus number of hits in NFB = 201 (chi-square=4.17, $p<.04$).

4. The (il)logic of the proposed 'natural' explanation

As Troscianko and Blackmore explain themselves there is a physical difference between the trials in the FB and the trials in the NFB condition. This difference is reflected in the program that controls the experiment and which can be represented by the following sequence:

1. RNG sample
2. if sample equals target then give FB
3. RNG sample
4. if sample equals target then don't give FB
5. prerecorded sample
6. if sample equals target then give FB
7. prerecorded sample
8. if sample equals target then don't give FB
9. return to 1.

The sequence is repeated 512 times and the target has been generated before the run.

It is obvious that prior to the third step a Feedback event might occur which might have some effect on the subsequent sampling process. This is what Troscianko and Blackmore describe as . . "the output of the RNG might be influenced by the thermal state of the computer and by input/output activity immediately prior to number generation" . . This does not hold for the samples in the feedback condition (step 1), since no input/output activity takes place prior to these trials. Of course this does not hold for the trials in the prerecorded condition since these have been sampled before the run and with no differences between the state of the computer preceding the samples at all. This latter point is used by Troscianko and Blackmore to justify their selection of only the RNG samples from 1. and 3..

The logical prediction according to the proposed mechanism would have been: there will be a non-random effect in the Instantaneously generated trials in the non-feedback condition and there will be no-effect (chance results) in the other conditions. This is exactly what they found but it does not account for the experimental data.

The experimental data show, on the contrary, an effect in the feedback condition and chance results in the non-feedback condition. The experimental percentual above-chance scoring in the 'instantaneously' generated and the 'prerecorded' trials in the feedback condition is both substantial (39% and 20% respectively) while the non-feedback conditions show chance scoring.

Troscianko and Blackmore are of the opinion (private communication) that the results of the actual experiment might be explained by assuming an extra mechanism that might be responsible for a general above chance scoring of the RNG for all conditions while still the differential effect would be produced by the change in input/output conditions. However in their control runs this proposed general above scoring inducing mechanism apparently has disappeared.

But I am even willing to follow their suggestion that there might have been some unspecified mechanism that disappeared mysteriously from experimental to control runs. The point is that although their model for the explanation of a differential effect is logical it just can not explain the magnitude of the differential effect observed in the experimental data. Actually any explanation based upon a potential interference between feedback-event and the subsequent sampling of the RNG can be ruled out. This can easily be seen by worst case analysis.

5. Worst case analysis

On the average the number of feedback events (just looking at the 'instantaneous' conditions) is 8 per run (of 2048 trials). Now we assume that the RNG, just after these events, will show the worst possible distribution of numbers. That is, it would produce always the very same number instead of a nice random number. Then we have to distinguish 2 cases:

a. This particular number is equal to the target. The expected number of hits in the NFB condition will be 16 in that case.

b. The number is unequal to the target. This means that the run is artificially shortened by 8 trials. The expected number of hits therefore becomes 8 minus 8/256. (p-hit = 1/256)

The first case is rather improbable (about 1/256) and indeed no such scores of around 16/run have been observed. Therefore we have to focus on the second case. But in this worst case the artificial differential effect accumulated over 33 runs would be about 1 hit! The actual difference however is 47 hits. Hence Troscianko and Blackmore are capable to explain at most 2% of the observed effect with the mechanism that they suggest.

6. What constitutes a control in psi-research?

Does this mean that the conclusion must be that psi is responsible for the observed data. Of course not. As I stated in the original article: ".. the building up of knowledge is a slow process that requires data of many similar experiments ..". I am not prepared to conclude anything on the basis of a single experiment. This holds for single experiments as well as for single controls. If one seeks the security of the 100% sure conclusion the empirical science is the wrong address.

However, the issue at hand raises a real problem which, I think, the field of parapsychology has to face sooner or later. This is the problem of what constitutes a control in psi research. There seems to be no proper definition, unless one assumes a theoretical model which specifies some constraints. This stresses again the need for theoretical models. I would strongly advocate publication selection criteria that favor work which is based upon such models (at the cost of selection criteria which favor publication of work with (statistically) 'significant' outcomes).

Dick J. Bierman
R.I.P.P.,
Alexanderkade 1,
1018 CH Amsterdam
The Netherlands

Call for Papers for the 29th Annual Convention of the Parapsychological Association

The 29th annual convention of the Parapsychological Association will be held Tuesday, August 5 to Saturday, August 9, 1986 at Sonoma State University in Rohnert Park, California. Persons interested in attending the convention may write to the Arrangements Chairperson, Rob Quider, JFK University, 12 Altarinda Road, Orinda, CA 94563, U.S.A..

Please read this call carefully; there are changes of instructions from previous years, intended to facilitate archival publication. All submissions must be carefully edited and finished in camera ready form for inclusion in the Convention Proceedings.

Anyone may submit a paper or a poster for consideration by the Program Committee. Papers may be empirical, theoretical, or methodological in nature, but the Program Committee will not consider papers published elsewhere prior to the convention.

Papers should adhere to the style of the 'Publication Manual of the American Psychological Association' (3rd edition). They should be typed on one side of 8.5 x 11 or A4 paper, and each page should have minimum margins of one (1) inch on all sides. The first sheet should have a centered title, author(s), and affiliation, followed by an abstract of no more than 300 words. Text should not exceed 12 single-spaced pages (6000 words), with no more than five additional pages for essential figures, tables, and references. If possible (e.g., using a word-processor) prepare and submit one single-spaced copy and three double-spaced copies, otherwise send four single-spaced copies. Four copies of proposed papers must be received by the Program Chairman by the deadline of April 31, 1986.

Presentation time will be 20 to 30 minutes, including a question period. Indicate exactly what audio-visual aid you will need, and indicate which of multiple authors will make the presentation. In absentia presentations will be allowed only in exceptional circumstances.

Posters are brief papers or other materials presented on poster board in an installation separate from the convention floor. Proposals for posters must include four copies of all material to be presented in the poster and the size of the required posters. Photocopies of photographs are acceptable. This material must reach the Program Chairman by April 31, 1986.

Members and associates of the Parapsychological Association may propose symposia, panel discussions and workshops. Symposia are formal presentations by participants on related topics. Proposals for symposia must include four copies of a summary sheet indicating title, chairperson, participants, order of presentation, and proposed time allotments, up to a total of 90 minutes, including discussion periods. Proposals must also include a full paper, prepared as detailed above, from each participant. This complete package must reach the Program Chairman by April 31, 1986.

Panel Discussions are informal round-table discussions intended to maximize spontaneous interactions of participants and the audience. Formal presentations should not exceed five minutes. Proposals for panel discussions must include four copies of a summary sheet including a title , chairperson, participants, order of presentation, and time allotments up to a total of 90 minutes, and an abstract of less than 500 words from each panelist. The complete package must reach the Program Chairman by April 31, 1986.

Workshops are informal presentations, discussions, or demonstrations. Proposals for workshops should provide a title, chairperson, participants, and descriptions of workshop activity. Workshops will not be listed as part of the formal convention program, but will be announced during the convention. Workshop proposals will be accepted until June 15, 1986.

Address all correspondence regarding the program to:

Dr. Roger Nelson, Chairman
Parapsychological Association Program Committee
School of Engineering/Applied Science
Princeton University
Princeton, NJ 08544
U.S.A.

EUR.J.PARAPSYCHOL., V.6, 1986

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART IV. THE SECOND GROUP OF EXPERIMENTAL SERIES

Hendrik G. Boerenkamp
University of Utrecht

In a previous paper the design of an experimental study of paranormal impressions of psychics was described (Boerenkamp, 1985a). Psychics are defined as persons who believe themselves able to obtain paranormal impressions at will. Usually, psychics are consulted by clients about problems related to themselves and to persons in their environment. The complete study consists of a predetermined number of series of sessions with a group of psychics and non-psychics in which the subjects gave their impressions about persons unknown to them. The series with the psychics were divided in three standard series, plus two subgroups of five experimental series each. Two of the three standard series were held at the start of the investigation (standard series A and B). These standard series were included to provide a description of the content of sessions when psychics apply their assumed psi abilities under conditions which resemble as much as possible the daily circumstances of sessions with clients (see Boerenkamp, 1985b, for the results). A further aim of including these series was to compare the content of these sessions with the content of the sessions of each of the experimental series, which were held to study the effect of a number of variables on the content of the sessions, i.e. the statements of the psychics about target persons. The experimental series were divided in two subgroups of five series each. The second subgroup was carried out when the first subgroup was completed and was introduced to the psychics as being more

'experimental' compared to the series in the first subgroup. The results of the first subgroup of five series (series 1-5) were discussed in Boerenkamp, 1985c.
In this paper the results of the second subgroup of experimental series (series 6-10) are discussed and compared with the results of the two standard series.

In the standard series the psychics were presented a photograph and an object belonging to a target person. Target persons were chosen from the environment of the sitter. The psychic was then invited to give paranormal impressions concerning this person. The psychic was aware that the sitter was acquainted with the target person, but he or she was not informed beforehand about special problems in the life of the target person. The psychic received immediate feedback to his statements in the form of an affirmation or a denial (confined informative action), occasionally followed by some clarification (extended informative action), which provided additional related information. No information unrelated to the topic being discussed was provided in the feedback.

Series 6,7,8 and 9 were presented and administered as a related set of sessions. In the sessions of these series the information available to the psychic was minimal compared to the sessions of the standard series. In none of these four series was feedback given. In each of the series a necklace was used as an inductor.
In the session of series 10 the information available to the psychic was maximal compared to the sessions of the standard series. The extra information in this series consisted of the physical presence of the target person.

The second subgroup of five experimental series were included to study the effect of the following variables:

(1) The inductor in non-feedback conditions (series 6 and 7).

As said, the series 6,7,8 and 9 were presented and administered as a related set of sessions comprising an 'experiment' by the two researchers. When the first necklace was presented the psychic was told that the necklace was one out of a set of four and that the researchers did not know the owner of the necklace. Two quite different necklaces used in series 6 and 7 belonged to the same target person, but the psychic was not informed about this. The experimenters did not know who the target person was. This was accomplished by requesting a third person to seek cooperation from a female friend and obtain from that target person two different necklaces , an expensive one and a cheap one. The necklace utilized as an inductor in series 6 was of silver. In series 7 the inductor was a simple necklace consisting of jagged, dull beads. By comparing series 6 and 7 the effect of different types of inductors in non-feedback conditions could be studied.

(2) The existence of a target person (series 7 and 8).

The two necklaces used as inductors in series 8 and 9 did not belong to a specific owner. This was accomplished by asking a third person to buy two simple necklaces which resembled the simple necklace used as an inductor in series 7. One of those had jagged, smooth beads (series 8) and the other one round, dull beads (series 9). Assumed paranormal impressions of the psychic in these series could not concern a particular person except perhaps the buyer. Because the necklaces used in series 7 and 8 were practically the same, it was possible to evaluate the effect of the existence or non-existence of a target person on the psychic's impressions.

(3) The social context (series 8 and 9).

In series 9 the psychic was requested to give his or her impressions in the absence of the investigators and to record them on tape. Since in series 8 and 9 no target person was involved and the inductors hardly differed, a comparison between the two series might indicate the extent to which the presence of sitter and observer influences the number and content of statements by psychics.

(4) The presence of the target person in the session (series 10).

In series 10 each of three target persons, aged about 25, 45, and 65 visited a subgroup of psychics accompanied by one of the experimenters. The conditions for this series were the same as those in the standard series except that in these sessions the psychic adressed the target person directly. The target persons were instructed not to indicate beforehand special problems and to avoid providing information unrelated to the topic being discussed, in the feedback.

The data from the sessions are described in terms of the number of statements by the psychic and the number of informative actions (feedback) by the sitter. Also the results of the informational, structural and interactional analyses are presented. In the informational analysis it is established which of the psychic's statements are sufficiently specific and spontaneous to assign them potential paranormal value, taking into account the available information at the moment the statement was made and whether those potential paranormal statements are correct. Such statements are called statements with positive paranormal value. In the structural analysis various characteristics of the set of statements are studied, e.g. the topics discussed in the statements. The interactional analysis concerns the actions psychics take when the sitter denies the correctness of the content of a statement, e.g. giving another interpretation of the impression. For a more detailed discussion of these concepts, see Boerenkamp, 1985a.

The two standard series A and B were carried out with 12 psychics. The experimental series 6,7,8 and 9 were conducted with 9 of the 12. Series 10 was carried out with 8 psychics. From the previous study (Boerenkamp, 1985b) it appeared that the behaviour of the psychics in standard series A was very consistent with their behaviour in standard series B. Hence, the data of standard series A and B were first combined in order to minimize the role of accidental influences on the content of the standard series. The combined data of the two standard series are compared with the data of each of the experimental series. (The combined data of standard series A and B are further denoted as standard series A+B). The comparison between the data of standard series A+B and the data of each of the experimental series is only based on the data of the psychics involved in both series. Thus the data of P3,P8 and P12 were excluded from the standard series A+B in the comparison between standard series A+B and series 6,7,8 and 9, the data of P8,P9,P11 and P12 were excluded from the standard series A+B in the comparison between standard series A+B and series 10.

Number of statements and informative actions

The number of statements made by the psychics in the different series is presented in table 1.

In each of the series 6, 7 and 8 the psychics made about the same number of statements. Thus, neither the type of necklace nor the

TABLE 1
Number of statements by the psychics

	A+B	6	7	8	9	10
P1	295	40	22	35	0	96
P2	129	22	40	18	0	104
P3	219	--	--	--	--	150
P4	147	5	9	6	2	101
P5	197	28	24	26	34	142
P6	213	47	30	22	78	243
P7	184	18	16	20	13	157
P8	81	--	--	--	--	--
P9	134	28	22	22	20	--
P10	259	32	23	40	0	318
P11	137	26	32	26	8	--
P12	126	--	--	--	--	--
all	2121	246	218	215	155	1311

existence of the target person influenced the number of statements. In standard series A+B (photograph and object as inductor and extended feedback from the sitter) the psychics (minus P3, P8 and P12) made 1695 statements. Thus, retaining feedback plus providing only an object as the inductor reduced the number of statements to a large extent (a mean number of 226 in each of the series 6, 7 and 8, a number of 848 in a comparable standard series).
In series 9, in which a social context was also retained, the number of statements further decreased to 155. However, the reactions of the different psychics to the request to give impressions in the absence of a sitter or observer were rather different. Three psychics accepted but did not succeed in the task of giving impressions in the absence of the investigators. For another psychic (P6) this highly 'experimental' task was a 'challenge'. She made significantly more statements in this session of series 9 than she did in the comparable session of series 8.

Further, it appeared that the psychics were consistent in the number of statements they made in the two standard series A and B. Each of them had a consistent preference for the duration of a session. The correlation between the length of the sessions of each psychic in standard series A+B (minus P3, P8 and P12) on the one hand and in the experimental series 6+7+8+9 on the other hand is marginally significant (Spearman $r= .61$, $df= 7$, $p< .05$). (All correlations applied in this paper are Spearman r correlations and are denoted further with rs; table P of Siegel's Nonparametric Statistics, McGraw-Hill, 1956, is consulted to find the significance level of the observed correlation). This implies that the administration of a related set of sessions comprising an 'experiment' in which psychics get a minimum of information hardly influences their individual preference for the duration of sessions.

In series 10 (target person present) the psychics made 1311 statements; the number of informative actions by the sitter was 1033 (344 extended and 689 confined ones). In standard series A+B the psychics (minus P8, P9, P11 and P12) made 1643 statements: the number of informative actions was 1055 (438 extended and 617 confined ones). Thus, it appeared that the sitters in series 10 (target persons who gave feedback about themselves) reacted more often than the sitters in the standard series (investigators who gave feedback about persons from their environment) did (chi-square= 12.55, $df= 1$, $p< .001$). However, they reacted only more often with confined feedback responses. The proportion of extended feedback responses is about equal in standard series A+B and in series 10 (344 extended responses to 1311 statements in series 10 and 431 extended responses to 1643 statements in standard series A+B). Hence, it appeared that the psychics made more statements when they communicated directly with the target person (1311 in series 10 and a number of 822 in a comparable standard series), but this considerable difference is probably partly due to the higher number of confined feedback reactions from the target person.

As was the case in standard series A+B the number of informative actions by the sitter depended on the number of statements by the psychic in series 10 ($rs= .98$, $df= 6$, $p< .01$).

It seems that the psychics were not consistent in the number of statements made in standard series A+B and in series 10 ($rs= .19$, $df= 6$, n.s.). This would imply that the presence of the target person would have a significant impact on their individual preference for the duration of a session. However, this is probably not the case. The non-significance of the correlation was caused by an 'artefact'. P1, who of all psychics made most statements in the sessions of standard series A+B had to stop the session in series 10 unvoluntarily because of an external reason (unexpected visitor) and made therefore the lowest number of statements in series 10. If the data of P1 are neglected the correlation is marginally significant ($r_s = .79$, $df = 5$, $p < .05$). This implies that the presence of the target person hardly influences their individual preference for the duration of a session.

The informational analysis

The procedure of selecting statements with potential paranomal value is described in detail by Boerenkamp (1984). The selection was based on the estimate of the probability of correspondence (specific versus vague) combined with the estimate of the degree of spontaneity (spontaneous versus inferred) of each statement by two judges. Thus, each statement was rated by two judges in two different ways on a four-point scale. On these scales a low degree of potential paranormal value is represented by a score of 1 or 2 and a high degree of potential paranormal value by a score of 3 or 4. In the study mentioned above it was found that two judges agreed on 82% of the statements to which of the categories (1 or 2 versus 3 or 4) it should be assigned.

The inter-rater-reliability for two judges in each of the five experimental series was comparable to the inter-rater-reliability obtained in the previous study. In series 6 the judges agreed on 77% of the statements, in series 7 on 76%, in series 8 on 79%, in series 9 on 78% and in series 10 on 81% of the statements. The distributions of the scores of potential paranormal value, based on the combined scores of two judges rating each statement on both probability of correspondence and degree of spontaneity on scale ranges from 1 to 4, are presented in table 2.
(In the rest of this paper the abbreviation ppv will be used for potential paranormal value).

According to the Kolmogorov-Smirnov two-sample test the distributions of scores of ppv in series 6 and 7, series 7 and 8, and series 8 and 9 do not significantly differ (e.g. series 6 and 7: $D_{max} = .07$, chi-square= 2.54, $df=2$, n.s.). (The Kolmogorov-Smirnov two-sample test is denoted further in this paper as KS). The distributions of scores of ppv in each of the series 6, 7, 8 and 9 significantly differ from the distribution of scores of potential paranormal value in the

TABLE 2
Distribution of scores of ppv in the
standard series A+B and the second five experimental series

score	low 4-7	(%)	low-med 7-9	(%)	medium 9-11	(%)	med-hig 11-13	(%)	high 13-16	(%)	total
series A+B	1174	(55)	479	(23)	280	(13)	119	(6)	69	(3)	2121
series 6	55	(22)	73	(30)	58	(24)	36	(15)	24	(10)	246
series 7	53	(24)	76	(35)	52	(24)	25	(11)	12	(6)	218
series 8	56	(26)	61	(28)	55	(26)	28	(13)	15	(7)	215
series 9	42	(27)	39	(25)	36	(23)	23	(15)	15	(10)	155
series 10	647	(49)	340	(26)	199	(15)	84	(6)	41	(3)	1311

standard series A+B (minus P3,P8 and P12) (e.g. series 6 and series A+B: KS: Dmax= .32, chi-square= 86.58, df= 2, $p<.0001$). The statements in each of these series were rated as having higher ppv.
The distribution of scores in series 10 does not differ significantly from the distribution of scores in the standard series A+B (minus P8, P9, P11 and P12)(KS: Dmax= .03, chi-square= 2.40, df=2, n.s.).
Applying a cut-off criterion between the medium and medium-high categories, each of the series 6, 7, 8 and 9 yielded about 21% (range 17% to 25%) statements with potential paranormal value, about twice as much as a standard series. Series 10 yielded 9% statements with ppv, the same proportion as the standard series.

Of the 188 statements of standard series A+B having ppv, 105 were made in the first half and 83 were made in the second half of the session. This difference was not significant. Comparable results were observed in each of series 6, 7, 8 and 9. In each series more statements with ppv were made in the first half of the session. However, the difference within each series was not significant. The preponderance of statements with ppv in the first half of the session is significant when the four series are combined (107 in the first half and 71 in the second half; chi-square= 8.75,df=1, $p< .01$).
Of the 125 statements of series 10 having ppv, 74 were made in the

first half and 51 in the second half (chi-square= 4.24,df=1, $p < .05$).

As the number of statements with ppv from each psychic in each of the series 6, 7, 8 and 9 is relatively low, the figures of the four series were first combined in order to study whether the psychics were consistent in the percentage statements with ppv in these series, compared with the standard series. The percentages of statements with ppv of each psychic for the different series are presented in table 3.

TABLE 3
Proportions of statements with ppv

series	A+B	6+7+8+9	10
P1	13%	13%	22%
P2	5%	15%	4%
P3	4%	---	3%
P4	4%	14%	2%
P5	5%	10%	13%
P6	9%	25%	6%
P7	21%	48%	17%
P8	11%	---	---
P9	3%	1%	---
P10	12%	27%	10%
P11	4%	39%	---
P12	11%	---	---
all	9%	21%	9%

The correlation between the proportion of statements with ppv over psychics in series A+B and series 6+7+8+9 is positive but not to a significant degree (rs= .46, df= 7, n.s.). The same applies for the correlation between the proportions in series A+B on the one hand and series 10 on the other hand (rs= .54, df= 6, n.s.).
These results suggest that as regards the proportion of statements with ppv, the psychics were influenced by a situation in which

feedback as well as a photograph of the target person were retained as well as by a situation in which the target person was present. Some psychics take more 'risks' by giving comparatively more spontaneous and specific statements in such situations compared to a situation in which they receive extended feedback, a photograph and an object of the target person. With other psychics an opposite effect is noticeable. However it has to be noted that despite the non-significance, the correlations are positive. This suggests that as regards the proportion of statements with ppv the impact of the feedback and the photograph on the one hand and the presence of the target person on the other hand is restricted.
The proportion of statements with ppv summarized over the series varied over psychics from 1% to 48% in the series 6+7+8+9 and from 2% to 22% in series 10. Hence between psychics there is variation in this respect and consequently one can make a distinction between more and less 'risk-taking' psychics.

The second step in the informational analysis was to establish how many of the about 21% (series 6,7,8 and 9) and of the about 9% (series 10) of statements having ppv also fulfilled the criterion of 'sufficient degree of correspondence'. The statements having ppv were rated on a four point scale as being true or untrue. Statements which received a score of either 1 or 2 were called statements of negative paranormal value (untrue statements) and statements which received a score of 3 or 4 were called statements of positive paranormal value (true statements).
Although the two necklaces used as inductors in series 8 and 9 did not belong to a specific owner, the statements in series 8 and 9 'theoretically' could have positive paranormal value. The statements of the psychics in these series concern perhaps the buyer and were applied to her. From table 4 it appears that in series 6,7,8 and 9, and in series 10 the degree of correspondence between the content of the statements and the facts about the person is low.

No difference was observed between the distributions of positive and negative paranormal statements in series 6 and 7, series 7 and 8, series 8 and 9 (e.g. series 6 and 7: KS: Dmax= 0.03, df= 2, n.s.). It further appeared that the distributions in each of the experimental series 6,7,8 and 9 differed significantly from the distribution in standard series A+B (minus P3, P8, and P12) (e.g. standard series A+B and series 6: KS: Dmax= 0.28, chi-square= 13.77, df= 2, $p < .01$). In all series a higher number of statements with ppv were untrue than in the standard series.

TABLE 4
Distribution of scores of 'degree of correspondence'

score	negative value				positive value			
	1	(%)	2	(%)	3	(%)	4	(%)
series A+B	101	(54)	60	(32)	23	(12)	4	(2)
series 6	49	(82)	9	(15)	2	(3)	0	(0)
series 7	31	(84)	4	(11)	1	(3)	1	(3)
series 8	34	(79)	6	(14)	2	(5)	1	(2)
series 9	31	(82)	6	(16)	1	(3)	0	(0)
series 10	68	(54)	38	(30)	12	(10)	7	(6)

No difference was observed between the distributions in standard series A+B (minus P8, P9, P11 and P12) and in series 10 (KS: Dmax= 0.06, chi-square= 1.16, df= 2, n.s.).

The structural analysis

Series 6, 7, 8 and 9 were introduced as being an 'experiment'. In the structural analysis it was first analyzed whether interesting differences appeared between the distributions in series 6 and 7, in series 7 and 8, and in series 8 and 9, with respect to each of the characteristics. As only two interesting differences appeared with respect to the different characteristics between the series (differences between series 8 and 9), the figures of the four series were first summarized with respect to each characteristic and then compared with standard series A+B. The two interesting differences will be discussed separately.

In each of the series 6,7,8 and 9, and series 10 the number of statements with positive paranormal value was too low to render a meaningful statistical comparison between distributions of statements with positive paranormal value and statements with negative paranormal value as regards the different characteristics studied. But there was no indication that the statements with positive paranormal value were different with respect to any of the characteristics. Therefore, in the structural analysis of series 6, 7, 8 and 9 and of series 10 only the set of statements with ppv and the set of statements without ppv were compared with respect to each characteristic.

(1) Topics discussed in the statements

The distribution of the topics in the standard series A+B, serie 6, 7, 8 and 9, and in series 10 is presented in table 5. The distributions of all statements (all) and statements having ppv are presented in real numbers for the main categories (A-E), whereas for the subcategories (10-54) percentages are given. (For some categories the description is abbreviated because of the limited space in the table).

The distributions of the topics were rather similar for the different series. The correlation between the frequencies of the topics (subcategories 10-45) for series 6+7+8+9 versus standard series A+B (minus P3, P8, P12) is marginally significant (rs= .57, t= 2.62, df= 14, p< .05). However, it has to be noted that the psychics made relatively more statements about physical characteristics in the 'experimental' series 6,7,8 and 9 than they did in the standard series.
The correlation between the frequencies of the topics for series 10 versus standard series A+B (minus P8, P9, P11 and P12) is significant (rs= .81, t= 5.16, df= 14, p< .001). However, it has to be noted that the psychics made relatively more statements about specific topics in series 10 than they did in the standard series.
Since in the sessions of series 6 and 7 the same target person was involved, it is of interest to study whether the content of the statements about that target person was consistent over the two sessions. It turned out methodologically impossible to classify the two sets of statements of each psychic as consistent or inconsistent with each other. Although each psychic discussed the different topics in each session about the same number of times, the content of the topic most often was different. For example, in the session of series 6 the psychic states that the target person has problems with her legs and in the session of series 7 the psychic states that the target person has diabetes. If only the two sets of statements with ppv of each psychic were compared neither a consistent nor an inconsistent pattern appeared. However, while hardly any identical statement having

TABLE 5
Distribution of topics discussed in the statements

	A+B	6+7+8+9	10
Total number of topics	2856	1169	1875
A Physical characteristics: all	462	397	402
Physical characteristics: ppv	73	99	64
10 Sex	0%	3%	0%
11 Age	1%	4%	2%
12 Appearance	5%	10%	5%
13 Bodily health	8%	11%	11%
14 Being alive or dead	2%	6%	3%
B Psychological charact.: all	1224	392	521
Psychological charact.: ppv	45	33	22
21 Personality traits	27%	15%	16%
22 Psychological circumstances	14%	15%	11%
23 Religious orientation	2%	3%	1%
C Relations: all	521	73	230
Relations: ppv	26	11	7
31 Relations with family members	8%	3%	6%
32 Relations with friends	7%	3%	6%
33 Relation with sitter (note)	3%	2%	0%
D Specific topics: all	649	307	722
Specific topics: ppv	93	112	83
41 Civil status	3%	3%	3%
42 Circumstances in work	9%	5%	9%
43 Circumstances in living	3%	5%	7%
44 Leisure activity	4%	3%	11%
45 Specific name, event	4%	11%	9%

Note: category 33: in series 6+7+8+9 relation with third person.

ppv was made in the two series, there were some clear inconsistencies. For example, a psychic said about the target person in the session of series 6 "Her name starts with a B" and in the session of series 7 "Her name starts with an A".
It has further to be noted that none of the psychics noticed that he or she was talking about the same person a second time.

In standard series A+B the distribution of topics (main categories A-D) discussed in the statements with ppv differed strongly from the distribution of topics discussed in the statements without ppv. In the set of statements with ppv fewer descriptions of psychological characteristics and relations with other people were observed. The same pattern was also observed in series 6, 7, 8 and 9 and in series 10.

(2) Person discussed in the topics

Out of the 1169 topics discussed in the statements of series 6+7+8+9, a total of 314 (27%) concerned a person related to the target person. In standard series A+B (minus P3, P8 and P12) 327 of the 2282 topics (14%) concerned a person related to the target person. The distributions in standard series A+B and series 6+7+8+9 significantly differ (chi-square= 79.43, df= 1, p< .0001). Thus, an object in non-feedback conditions, induced more impressions about persons related to the target person. The preponderance of statements about persons related to the target person is first of all the result of more statements about dead relatives of the target person (see category 14 in table 5).
Out of the 1875 topics discussed in the statements of series 10, a total of 518 (28%) concerned persons related to the target person. The comparable data in standard series A+B (minus P8, P9, P11 and P12) were 314 out of 2217 topics (14%). Thus, the number of topics about persons related to the target person is also higher in the case in which the target person is present (chi-square= 112.85, df= 1, p< .0001).

Further, it appeared in standard series A+B that statements with ppv involved relatively more often topics concerning a person related to the target person (26%) than statements with no ppv (11%). Apparently when psychics talk about persons related to the target person they are for example less likely to concentrate on the psychological characteristics of these persons and more likely to discuss their appearance. In series 6+7+8+9 the set of statements with potential paranormal value involved the same number of topics concerning a person related to the target person than statements without potential paranormal value (chi-square= 3.01, df=1, n.s.). Thus, the ppv of statements about the target person is equal to the ppv of statements about persons related to the target person in such 'experimental' conditions. In series 10 the pattern was in accordance with the pattern in the standard series (chi-square= 31.87, df=1, p< .0001).

(3) Number of statements about past, present and future

For the different series the distributions of statements about past, present and future are presented in table 6.

TABLE 6
Distribution of statements about past, present and future

	past	past %	present	present %	future	future %
series A+B	339	16%	1584	75%	198	9%
series 6+7+8+9	228	27%	539	65%	67	8%
series 10	242	18%	734	56%	335	26%

The distributions of statements about past, present and future in standard series A+B (minus P3, P8 and P12) and series 6+7+8+9 are significantly different (chi-square= 37.48, df= 2, p< .0001). The psychics made more statements about the past in the 'experimental' series than in the standard series. The predominance of statements about the past is first of all the result of more statements about dead relatives of the target person in these four series (see category 14 in table 5).
The distributions in series A+B (minus P8, P9, P11 and P12) and series 10 are also significantly different (chi-square= 153.39, df= 2, p< .0001). If the target person is present the psychics make much more statements about the future. This result suggests that target persons in the daily practice of psychics show first of all interest in statements about the future.

When one splits up the sessions into first and second halves in order to see whether the session as a whole has a temporal pattern, it appeared that there was a predominance of statements about the future in the second half of the sessions in the standard series A+B. The distributions of statements about past, present and future in the first and second half of the sessions for the different series are presented in table 7.

TABLE 7
Distribution of statements about past, present and future in the first and second halves of the sessions

	past		present		future	
	FH	SH	FH	SH	FH	SH
series A+B	162	177	828	756	71	127
series 6+7+8+9	111	117	285	254	22	45
series 10	152	90	366	368	138	197

Note: FH: first half; SH: second half.

It appeared that the predominance of statements about the future in the second half of the sessions was consistent in series 6+7+8+9 (chi-square= 9.83, df= 2, $p < .01$). In series 10 statements about the past were predominant in the first halves of the sessions and statements about the future were predominant in the second halves of the sessions (chi-square= 26.28, df= 2, $p < .0001$).

Further, in standard series A+B it appeared that the set of statements with potential paranormal value contained relatively more statements about the past (29%) and relatively fewer statements about the present and the future compared to the set of statements without ppv (15%). This difference was also consistent over the different experimental series (series 6+7+8+9: 34% versus 27%; series 10: 23% versus 18%).

(4) Number of statements about a favourable, neutral or unfavourable state of affairs.

In the comparison between the two standard series A and B the only strong inconsistency between the two series appeared to be the difference in the number of statements about favourable and unfavourable states of affairs concerning the target person and related persons (Boerenkamp, 1985b). This number was influenced to a certain degree by the different ways the psychics perceived the target persons. For instance, in certain cases psychic 1 had a positive impression and psychic 2 a negative impression of the same person. Therefore, interpretations of differences in this respect between standard series A+B and each of the experimental series are tentative, since a part of the difference has to be attributed to this 'standard' variation. The distributions of statements about a favourable, neutral or unfavourable state of affairs are presented in table 8.

TABLE 8
Distribution of statements concerning a favourable, neutral or unfavourable state of affairs

	favourable	fav %	neutral	neu %	unfavourable	unf %
series A+B	419	20%	753	35%	949	45%
series 6+7+8+9	126	15%	376	45%	332	40%
series 10	243	19%	540	41%	528	40%

The distributions in standard series A+B (minus P3, P8 and P12) and series 6+7+8+9 are significantly different (chi-square= 13.63, df=2, $p < .01$). The difference between standard series A+B and series 6+7+8+9 appeared primarily to be the result of a difference in number of statements about a neutral state of affairs in series 6+7+8+9. Less statements about a favourable as well as less statements about an unfavourable state of affairs were made by the psychics when they only employed an object as an inductor in non-feedback conditions.

The distributions in standard series A+B (minus P8, P9, P11 and P12) and series 10 are different to a marginally significant degree (chi-square= 8.42, df= 2, $p < .02$). The psychics made somewhat less statements about an unfavourable state of affairs when the target person was present.

When one splits up the sessions of standard series A+B into first and second halves in order to see whether the sessions as a whole had an 'emotional' pattern, it appeared that statements about an unfavourable state of affairs slightly dominated in the first half of the session (chi-square= 5.89, df=2, $p < .05$). The distributions of statements in first and second halves of the sessions for the different series are presented in table 9.

TABLE 9
Distribution of statements concerning a favourable, neutral or unfavourable state of affairs
in the first and second halves of the sessions

	favourable		neutral		unfavourable	
	FH	SH	FH	SH	FH	SH
series A+B	198	221	361	392	502	447
series 6+7+8+9	66	60	192	184	160	172
series 10	115	128	271	269	270	258

Note: FH: first half; SH: second half.

It appeared that the slight predominance of statements about an unfavourable state of affairs in the first half of the sessions was not consistent over these series. Neither in series 6+7+8+9 (chi-square= 0.88, df=2, n.s.) nor in series 10 (chi-square= 0.97, df= 2, n.s.) a difference was observed. In Boerenkamp 1985b it was suggested that the slight predominance of statements about an unfavourable state of affairs in the first half of the sessions might indicate that psychics are accustomed to sitters who indicate problems at the beginning of a session. The target persons in series 10 were

instructed not to do so. Therefore, it seems that psychics distribute statements about favourable states of affairs and statements about unfavourable states of affairs equally over the session if the target person does not indicate problems beforehand.
The set of statements with potential paranormal value in standard series A+B had the same distribution with respect to this characteristic as the set of statements without ppv (chi-square= 2.09, df= 2, n.s.). This was not the case in series 6+7+8+9. In the set of statements with ppv were relatively more statements about a neutral state of affairs. In the set of statements with potential paranormal value in series 10 statements about a neutral state of affairs dominated to a marginally significant degree (chi-square= 7.17, df= 2, $p < .03$).

(5) Number of statements in the form of advice

In standard series A+B, 120 statements (6%) involved advice. In series 6+7+8+9 this number was 73 statements (9%) and in series 10, 194 statements (15%). The number of statements involving advice in standard series A+B (minus P3, P8 and P12) differed to a marginally significant degree from the number of statements involving advice in series 6+7+8+9 (chi-square= 4.94, df= 1, $p < .03$). Thus, the number of statements involving advice increased in the more 'experimental' series as compared to the standard series.
The number of statements involving advice in standard series A+B (minus P8, P9, P11 and P12) differed to a strongly significant degree from the number of statements involving advice in series 10 (chi-square= 66.21, df= 1, $p < .0001$). Thus, the number of statements involving advice increased in the presence of the target person.

It appeared that the statements with ppv contained fewer statements involving advice compared to the other statements in series 6+7+8+9 (6% versus 10%). However, this difference is not significant (chi-square= 2.30, df= 1, n.s.). The statements with ppv in series 10 contained significantly fewer statements involving advice compared to the other statements (6% versus 16%; chi-square= 8.48, df= 1, $p < .01$).

(6) Number of times a silence precedes a statement

In standard series A+B, 337 statements (16%) were preceded by a silence of 3 seconds or more. In series 6+7+8+9 this number was 306 statements (37%) and in series 10, 103 statements (8%).
The number of statements preceded by a silence in standard series A+B

(minus P3, P8 and P12) is significantly different from the number of such statements in series 6+7+8+9 (chi-square= 135.81, df= 1, $p < .0001$). The number of silences is more than twice as much as in a standard series. An interesting difference between series 8 and 9 is that the number of silences in series 9 (17%) was significantly lower than in series 8 (36%). Probably most psychics formulated their impressions before they recorded them on tape (see the introductional part of this article), causing fewer pauses.
The number of statements preceded by a silence in standard series A+B (minus P8, P9, P11 and P12) is significantly higher than the the number of such statements in series 10 (chi-square= 32.22, df= 1, $p < .0001$). It has to be remembered that the target persons in series 10 reacted more often with confined feedback reactions than the sitters in the standard series. Probably this contributed to a lower number of silences in series 10.

Further, it appeared that with respect to this characteristic the set of statements with potential paranormal value in standard series A+B was not significantly different from the set of statements without ppv. However, in series 6+7+8+9 more statements with potential paranormal value were preceded by a silence (chi-square= 8.06, df= 1, $p < .01$). Thus, in the condition in which only an object was presented and in which they received no feedback at all, the psychics tried more often to obtain spontaneous and specific impressions during the silences.
The sets of statements with and without ppv did not differ in this respect in series 10 (chi-square= 0.01, df= 1, n.s.).

(7) Number of positive and rhetorical statements

Rhetorical statements are statements in which the psychic asks for immediate feedback in contrast which positive statements in which the psychic does not ask for immediate feedback. The distributions of positive and rhetorical statements for the different series are presented in table 10.

It has to be remembered that in series 6+7+8+9 the psychic was informed that the sitter was not acquainted with the target person. Therefore, one would expect only positive statements in series 6+7+8+9. However, it appeared that psychics are so accustomed to receiving feedback that a considerable number of times (8%) they simply neglected the information from the sitter that he was not acquainted with the target person by making rhetorical statements.

TABLE 10
Distribution of positive and rhetorical statements

	positive	positive %	rhetorical	rhetorical %
series A+B	1474	69%	647	31%
series 6+7+8+9	769	92%	65	8%
series 10	852	65%	459	35%

An interesting difference between series 8 and 9 is that the number of rhetorical statements in series 9 (1%) was significantly lower than in series 8 (7%). If requested to give his or her impressions in the absence of a sitter the psychics hardly make rhetorical statements. The distributions of positive (non-rhetorical) and rhetorical statements in standard series A+B (minus P8, P9, P11 and P12) and series 10 are significantly different (chi-square= 11.03, df= 1, p< .001). If the target person is present the psychics make more rhetorical statements.

When the sessions of standard series A+B were split up into first and second halves, a significant difference appeared between the first and second halves of the sessions as regards the number of rhetorical statements. More rhetorical statements were found in the first half of the sessions (chi-square= 14.32, df= 1, p< .001). In table 11 the distributions of rhetorical and positive statements are presented for series A+B, series 6+7+8+9 and series 10.

The predominance of rhetorical statements in the first half of the sessions was not found in the 'experimental' series 6+7+8+9. In series 10 was a marginally significant difference between number of positive and rhetorical statements in the two parts of the session observed (chi-square= 4.26, df= 1, p< .05). Hence, in series 10 (target person present) the psychics tried to elicit more feedback in the first half of the session.

The set of statements with ppv in standard series A+B contained more

TABLE 11
Distribution of positive and rhetorical statements in the first and second halves of the sessions

	positive statements FH	positive statements SH	rhetorical statements FH	rhetorical statements SH
series A+B	700	780	361	280
series 6+7+8+9	383	386	34	31
series 10	408	444	248	211

Note: FH: first half; SH: second half.

rhetorical statements compared to the set of statements with no ppv (chi-square= 112.24, df= 1, $p < .0001$). Table 12 presents the distributions of positive and rhetorical statements with potential paranormal value (ppv) and with no potential paranormal value (nppv) for the different series.

TABLE 12
Distribution of positive and rhetorical statements with and with no potential paranormal value

	positive statements ppv	positive statements nppv	rhetorical statements ppv	rhetorical statements nppv
series A+B	67	1413	121	520
series 6+7+8+9	157	612	21	44
series 10	53	799	72	387

This finding in standard series A and B appeared to be consistent over the different series (chi-square= 4.37, df= 1, p< .04 in series 6+7+8+9 and chi-square= 29.90, df= 1, p< .001 in series 10).

(8) Number of statements preceding an informative action

Since all sorts of statements might invite feedback it was calculated how many statements were made by the psychic before the sitter reacted with an informative action. If the psychic made a statement and the sitter reacted directly with one or more informative actions, it is indicated by P1S in table 13. If the psychic made two statements before the sitter reacted, it is indicated by P2S etc..

TABLE 13
Distribution of statements preceding an informative action

	P1S%	P1S	P2S	P3S	P4S	P>4S
series A+B (c)	42%	695	213	67	27	35
series 10	66%	867	113	29	12	12

Note: (c): corrected for P8, P9, P11 and P12.

The distributions in series A+B (minus P 8, P9, P11 and P12) and series 10 are significantly different (chi-square= 81.67, df= 4, p< .001), which implies that the target persons gave more direct feedback reactions as the sitters did in the standard series.

In standard series A+B the set of statements with ppv contained on the average the same number of statements which were part of P1S interactions as the set of statements with no ppv (chi-square= 0.02, df= 1, n.s.). A similar result was found in series 10 (chi-square= 2.92, df= 1, n.s.).

The interactional analysis

The objective of this analysis is the actions taken by the psychics after receiving a denial as a reaction to one of their statements. For the standard series it was found that the psychics mainly use four types of responses to a denial:

(1) Accepting the denial by giving another impression (AD)
(2) Giving a new interpretation to the denied impression (NI)
(3) Suggesting the target person knows better than the sitter (T>S)
(4) Suggesting the content of the informative action is equal to the content of the statement (I=S)

In the discussion of the results of the standard series (Boerenkamp 1985b) it was suggested that it was conceivable that other categories of reaction to a denial had to be added to deal with the 'interactional' richer situation in series 10, in which the target person was present. This appeared not to be the case. If the target person is present the psychics do not use category 3. Instead they suggest that they know better as the target person himself (P>T).
For example: P: You must be careful of your neighbour
T: I don't think so, he is nice
P: That is what you think but you must be careful

The AD and NI reactions are called the 'acceptance of failure' reactions and the T>S (or P>T in series 10) and I=S reactions are called the 'resistance against failure' reactions. In table 14 the different responses of the psychics for the different series are presented.

The distribution in standard series A+B and series 10 are different to a marginally significant degree (chi-square= 7.96, df= 3, p< .05). In the case of the target person being present the psychics tend to react more often with suggesting that the content of the informative action is equal to the content of the statement (I=S).
The number of denials by the target person (series 10) was higher than the number of denials by the sitters in the standard series (chi-square= 22.82, df= 1, p< .001).

DISCUSSION

As regards the length of the sessions, it appears first of all that the number of statements from the psychics in a session is minimal in the most 'experimental' conditions (only an object of the target

TABLE 14
Distribution of types of responses to a denial

	AD	NI	T>S P>T	I=S	acceptance AD+NI %	resistance T>S + I=S % P>T
series A+B (c)	19	75	18	12	76%	24%
series 10	23	96	20	39	67%	33%

Notes: T>S in series A+B; P>T in series 10.
(c): corrected for P8,P9,P11 and P12.

person and no feedback about the target person) and maximal in the most 'usual' conditions (target person present and feedback from the target person) as compared to the standard conditions (photograph and object of the target person and feedback about the target person).
In addition, it appears that the existence or non-existence of a target person does not influence the number of statements and that the number of statements is rather unpredictable if the psychics are requested to give their impressions in the absence of a sitter and to record them on tape. Some psychics accept but do not succeed in such a task while other psychics may experience such a task as a challenge. In general, the length of such 'sessions' is minimal.
Further, it appears that as regards the number of statements, psychics are rather consistent over the different series. The different experimental manipulations did not have much effect on their individual preference for sessions of a certain length.

In the standard series it was found that approximately only one out of ten statements could be considered as meeting the criterion of 'being sufficiently spontaneous and specific', i.e. could be assigned potential paranormal value (ppv) and that approximately only one out of ten of these statements with ppv could be considered to meet the criterion of 'being sufficiently correct', i.e. could be assigned positive paranormal value. According to the author's estimation none of the statements which met both criteria could be considered to have

a probability lower than 1 in 100. Therefore, it was concluded that the observed 'unexplainable' correspondences between the statements and the facts concerning the target person could be satisfactorily explained by assuming chance coincidence and that it is not necessary to assume a specific paranormal ability of the psychics.
The data in each of the series of the second group of experimental series support this tentative conclusion. A significant difference was observed in the distributions of scores of ppv in the series 6+7+8+9 as compared with the standard series A+B. About twice as many statements were assigned ppv in the series 6+7+8+9 (about 20% versus about 10% in the standard series). However, also about twice as many statements with potential paranormal value from the series 6+7+8+9 as from the standard series A+B did not meet the criterion of 'being sufficiently correct'. Therefore, in congruence with the standard series only about 1% of the statements in series 6+7+8+9 met both criteria.

Thus, in sessions in which the information available to the psychic is minimal, the statements with potential paranormal value seem more based on elements from the memory and imagination of the psychic.
The results in series 10 (target person present) are in congruence with the results in the standard series (about one out of ten statements meets the criterion of 'being sufficiently spontaneous and specific' and about one out of ten of these statements meets the criterion of 'being sufficiently correct').
In a previous paper (Boerenkamp 1985c) it was concluded that it is very doubtful whether research with psychics offers a more promising method than the methods currently employed in parapsychology (controlled studies of selected or unselected subjects in the laboratory or in daily life) as none of the variables under study in the first subgroup of experimental series (type of inductor in feedback conditions, amount of feedback and acute importance of the problem of the target person) had a facilitating or inhibiting impact on the assumed ability. The data of the second subgroup of experimental series support this conclusion. Neither the 'experimental' conditions nor the presence of the target person had a facilitating or inhibiting effect.

As regards the type of inductor in non-feedback conditions, the existence or non-existence of a target person, and the presence or absence of a social context in which the psychic gives his or her impressions, the following findings result from the structural analysis (characteristics of the set of statements):

None of these variables has an significant impact on the characteristics studied except that the absence of a social context results in a lower number of statements preceded by a silence and a lower number of rhetorical statements. Probably, most psychics formulated their impressions before they recorded them on tape.

Concerning the combined data from series 6+7+8+9, presented as a related set of 'experimental' sessions in which the information available to the psychic was minimal, the following findings result from the structural analysis compared with the findings in the standard series:
(1) The number of statements about physical characteristics increases especially at the cost of statements about psychological characteristics. The psychics concentrate more on age, appearance, bodily health and being alive or dead.
(2) The number of statements about persons related to the target person increases.
(3) The number of statements about the past increases.
Both the predominance of statements about persons related to the target person and the predominance of statements about the past is first of all the result the result of more statements about dead relatives of the target person.
(4) The number of statements about a neutral state of affairs increases.
(5) The number of statements involving advice increases.
(6) The number of statements preceded by a silence increases.
(7) The number of rhetorical statements decreases.

As regards the presence of the target person the following findings appear from the structural analysis as compared to the findings in the standard series:
(1) The number of statements about specific topics increases especially at the cost of statements about psychological characteristics. Altough the target persons were instructed not to indicate a special topic beforehand, most psychics made a considerable number of statements about the leisure activity of the first target person, about circumstances in work of the second target person and about circumstances in living of the third target person. Probably they noticed a preference for those topics of each target person.
(2) The number of statements about persons related to the target person increases. Psychics made a considerable number of statements about friends, children or spouses of the target persons.
(3) The number of statements about the future increases, especially in

the second half of the session. Probably the target persons were most interested in statements about the future and probably the psychics noticed this preference.
(4) The number of statements about an unfavourable state of affairs decreases. Perhaps this was caused partly by the fact that the target persons were instructed not to indicate special problems. Several psychics explicitly asked 'what the problem was'.
(5) The number of statements involving advice increases.
(6) The number of statements preceded by a silence decreases.
(7) The number of rhetorical statements increases.
(8) The number of PIS interactions (the psychic makes a statement and the sitter reacts directly with one or more informative actions) increases.

As regards the set of statements having ppv compared to the set of statements with no ppv, most differences observed between the two sets of the standard series are present in these series as well. The most remarkable difference is the following: in the set of statements with potential paranormal value, relatively more statements were preceded by a silence in the 'experimental' series in which only an object was presented and the sitter provided no feedback than in the set of other statements. This difference was not found in the standard series. However, the same difference was also found in experimental series 1 (only an object of the target person and feedback) as well as in experimental series 4 (photograph and object of the target person and no feedback). Thus, the data from the 'experimental' series confirm the conclusion (Boerenkamp 1985c) that only in relatively difficult task-circumstances the psychics use the silences more often to receive more spontaneous and specific impressions. In circumstances in which the target person was present and which might be considered as the relatively easiest ones the sets of statements with and without potential paranormal value did not differ in this respect.

The results of the interactional analysis in the standard series indicated that psychics most often react to a denial by giving a new interpretation to their impression. They also did this in the presence of the target person. The presence of the target person had only a slight impact on this type of behaviour of the psychics. If the target person is not present the psychics sometimes suggest that the target person would affirm the denied statement. If the target person is present the psychics sometimes simply declare instead that they know better than the target person himself or herself.

ABSTRACT

The aim of the analyses presented in this paper was to study the effect of some variables on the content of sessions of psychics. These variables were (1) the type of inductor in non-feedback conditions, (2) the existence of a target person, (3) the social context in which the psychic gives his or her impressions, and (4) the presence of the target person.
It was found that the number of statements is directly related to the the amount of available information to the psychic. More available information gives more statements. Experimental manipulations hardly affect the individual preference of psychics for sessions of a certain length.

None of the variables studied had a significant influence on the number of statements with positive paranomal value, i.e. spontaneous and specific statements which are correct. Only about 1% of all statements seem to be appropriate, but that can be expected by the chance-hypothesis. In view of the judging procedures applied, the low percentage is not supportive of a paranormal interpretation. This result, combined with the results in the first subgroup of experimental series implies that research with psychics does not offer a more promising research method than the methods employed most often in parapsychology.

Both the conditions in which the psychics get minimal information (only an object of a target person and no feedback) as well as maximal information (target person present and feedback from the target person) have a considerable influence on the structure of the behaviour of psychics as compared to standard conditions (photograph and object of the target person and feedback from a person related to the target person). Among other things: if the psychics get minimal information the number of statements involving advice increases and the number of statements preceded by a silence increases; if the psychics get maximal information the number of statements about the future increases, the number of statements involving advice increases, the number of statements preceded by a silence decreases and the number of rhetorical statements increases.

REFERENCES

Boerenkamp, H.G., Schouten, S.A. 'Estimating the potential paranormal value of verbal statements', J.o.P., 47, 1983, 121-130.

Boerenkamp, H.G. 'Potential paranormal value of statements of psychics acquired under feedback conditions', E.J.P., 5, 2, 1984, 101-124.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part I. Experimental design', E.J.P., 5, 4, 1985a, 327-348.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part II. The standard series', E.J.P., 5, 4, 1985b, 349-371.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part III. The first group of experimental series', E.J.P., 6, 1, 1985c, 33-70.

EUR.J.PARAPSYCHOL., V.6, 1986

MINIMIZING SUBJECT FRAUD IN PARAPSYCHOLOGY LABORATORIES

Robert L. Morris
Psychology Department
University of Edinburgh

The reduction of fraud by subjects in parapsychological research may involve much more than occasional consultations with professional magicians. In what follows, the author presents a preliminary organisation of strategies for minimizing attempts at fraud based in part on a consideration of the context of such attempts. The emphasis is on prevention more than detection; ideally the potential fraud (or 'pseudopsychic') should be dissuaded from taking the trouble to organise an attempt in the first place.

If we are to dissuade pseudopsychics from plying their trade with a researcher or institution, it is important first to understand the several motivations for fraud. Once these are understood, the researcher can develop strategies for reducing the availability of such rewards. The researcher must then communicate with the pseudopsychics either directly or indirectly, such as to persuade them of two things; (1) that if they attempt fraud they will almost certainly fail or if successful, such success will be short-lived and they will be detected; and (2) even if they are undetected there will be little reward, so little that it does not justify the risks and effort involved.

However, such communication may not always take place appropriately and certainly should not be relied upon. Thus part of the overall strategy of fraud prevention is to develop and apply adequate safeguards in the research situation itself. In fact, the presence of such safeguards becomes part of the information to be communicated to pseudopsychics.

MOTIVATIONS OF PSEUDOPSYCHICS

Why should one want to simulate psychic abilities? One group of motivators is personal gain. This can be financial, as in the case of the pseudopsychic who charges large sums to provide counselling or constructs a plan to swindle a gullible investor. Personal gain can also take the form of personal fame. Some people enjoy being the centre of attention and will strive for public recognition even if no money is necessarily involved. A third area of gain is personal power. The Rev. Jim Jones of Jonestown notoriety is a strong example of this. He maintained almost total control over the lives of his parishioners and followers by the systematic faking of Biblical miracles. Power and control can be exerted at many levels; or it can be widespread, as in the case of the psychic who publishes a bestseller combining personal claims of psychic power and psychic predictions of what bad things will happen if a given country does not follow a particular political or religious course. Many pseudopsychics, however, may seek only to be known and exert control with respect to a small, well-defined group. At the extreme, one might fake psychic ability solely to impress a prospective mate or lover.

Closely related to personal gain is enhanced self-esteem, which can include the desire to be socially helpful. Some pseudopsychics may fake psychic ability in order to persuade people who are dissatisfied with existing health care delivery to come instead to them for counselling or treatment, thereby feeling that they have accomplished a social good. Included here may be people who regard themselves as genuinely psychic but who need to practice pseudopsi to enhance their reputations or maintain them during periods of poor psychic performance. Self esteem may also involve the thrill of meeting successful challenges; thus an aspiring young magician may take pride and pleasure in defeating the safeguards of a perceived scientific elite at a prestigious institution, whether or not he/she ever discusses such feats publicly.

The above motivations are generally best served by avoiding exposure. As parapsychology is well aware, however, some motivations can involve the perpetration of a hoax followed by deliberate self-exposure. Magicians seeking publicity and status for their skills as magicians may follow such a course, as may someone wishing to evaluate or discredit the competence of an individual researcher, research unit, or parapsychology as a whole.

THE MESSAGE TO THE PSEUDOPSYCHIC

Ideally, a pseudopsychic should come to be discouraged from attempting to perpetrate fraud with a given research group, by being persuaded of at least three things.

1. Pseudopsychics should be persuaded that there is a high likelihood that they will be detected. As a starting point, the research group should identify and avail themselves of relevant expertise. One component is the educated researcher. Individual researchers should learn the general concepts and strategies of the simulation of psychic skill. A solid background in experimental and cognitive psychology is also very helpful, to understand the ways we organise and interpret information from our environments. Elsewhere the present author attempted to develop a conceptually rooted organisation of such strategies, using the observer of apparent psychic events as the reference point. My scheme is now being expanded and modified considerably, and others may develop comparable schemes that will be more effective. Material on the simulation of psi has been incorporated in my own parapsychology coursework since 1974, and in more recent years in the Summer Institute of the FRNM, the Master's program at John F. Kennedy University, and in various articles and reports.

Researchers who need specialised knowledge due to the nature of the claims they are investigating may need to become conversant with the tactics in that area as well as strategy. Learning tactics would mean learning specific details of the psychology, physics, and (where appropriate) biology of routines developed to accomplish specific 'psychic' effects. This can be done by a mix of reading a specialised literature available almost exclusively from magic shops and mail order services, plus close observation of practitioners willing to give tutorials.

In addition to the development of personal understanding of the general concepts and relevant specific tactics involved, researchers can avail themselves of procedures and guidelines developed by others sharing similar areas of interest. Within parapsychology, groups of workers in separate institutions will have similar problems to solve, and it is important that generalisable solutions be articulated. Such solutions may involve experimental materials, procedures, monitoring systems, and so on. It is helpful to understand both solutions and the rationales behind the solutions. Many of the card-guessing procedures evolved at Duke, such as screened touch matching (STM), down through (DT) technique, and so on, were developed in part to guard against certain fraud tactics. Those who superficially learned the techniques without learning the full logic behind them would not necessarily be able, later on, to modify these techniques intelligently, without losing the safeguards they provided. Once, when I was just becoming interested in conjuring and mentalism, I asked a senior researcher to tell me all about blindfolds, as I'd heard he was an expert. His response: "Don't ever use them; that's all you need to know."

Even with diligence in familiarising oneself with available expertise, there may still be occasions in which the on-site advice of a consulting professional magician can be extremely helpful. It is important to note that magicians specialise. Expertise of one sort, e.g. conjuring or bizarre magic, may not necessarily be an indicator of expertise in mentalism or close up technique. The consultant should be chosen to match the job. The Parapsychological Association has placed notices in trade journals of professional magicians, and is now developing a listing of magicians with diverse expertise who are willing to consult with parapsychological researchers.

At present there are no good guidelines as to how such consultations should proceed, and there are several interesting issues. First, when should consultation take place? Tentatively, I recommend two circumstances. One is general, in that it would seem ideal to have a magician with broad competence to critique a lab's procedures once a year, or at any time that a major procedural overhaul is carried out, e.g. through the introduction of new equipment or a major new area of research. The other is specific, involving any time a special claim is investigated such that the claim calls for testing procedures that are not part of the laboratory's established protocols, especially if the claims involve the ostensible talents of one or two special individuals.
Second, how much contact should the magician have with the claimants?

Ideally, the magician should meet the claimants, discuss the claims with them, and evolve a mutually acceptable protocol (if none currently exists). The magician should not be seen as having an adversary relationship with the claimant; if the claims are legitimate, the magician is really working in everyone's best interest. If such a meeting is impractical, a magician can often rely upon a videotape to get some feeling for the modes of operation to be involved. Failing this, our magical consultant is left at best with a rather indirect representation of the claims to be researched. Since a pseudopsychic would probably change the nature of the claim to negotiate a weaker protocol, often such indirect interaction leaves much to be desired. In fact, one advantage of a direct interaction between consultant and claimant is that the claimants can be led to think about and specify their claims (and boundaries thereof) more clearly, such as to be more amenable to good experimental protocol.
Third, what actually can a magician accomplish? As mentioned above, they may be able to develop a clear picture of the claims being made, and devise a good test of them. Or, they may be able to critique an existing procedure. They may be able to observe a claimant's performance and evaluate the likelihood that trickery was used, in some cases spotting the exact procedure used. Once a procedure is devised, they may be able to supervise its conductance and either be present physically during testing or train the researchers in how to run the procedure and detect attempts to circumvent it. The magician may be able to modify the procedure to take into account observed weaknesses, claimant complaints, and so on.
Fourth, what are the limits to what the magician can do? No magician can guarantee a foolproof protocol, and none can guarantee that there was no break in the protocol during a given session by claimant or researcher. The latter is especially true if the magician is not physically present during the session. Magicians can and do fool each other in demonstrations, although the tactics for fooling another magician are often a bit different. People fool us in part by capitalising on our assumptions, and magicians make slightly different assumptions when observing a performance than do the rest of us. Thus by calling in a magician as consultant we do not guarantee a foolproof experiment; we merely decrease the likelihood that some sort of fraud will take place, we learn more about how to design procedures, and we provide a public indicator that we have availed ourselves of such expertise.

The above is all well and good, but there may come times when researchers are unable to access relevant expertise, especially when doing field research or confronting a surprise visitor to one's facilities. This leads us to a rather speculative and controversial kind of resource - the expert system, a set of programs accessible by computer, representing a collection and organisation of expert knowledge of some sort. Such systems are presently being developed within artificial intelligence (AI) to make available otherwise remote expertise, for such diverse endeavours as oil drilling and medical diagnosis. No such systems presently exist for parapsychology, but they could be developed.

As a first step a problem area would be delineated, in this case the problem of designing fair tests of a host of psychic claims. Then a small body of experts would be selected, representing a mix of magicians, psychologists, physicists and parapsychologists, carefully selected. These experts would then be interviewed by knowledge engineers, whose job would be to extract, enhance, organise and represent the collective expertise in such a way that it could be assimilated by computer software. The result would be a complex program, or set of programs which would provide a potential user with access to a representation of that expertise. The result could take several forms; typical would be a query system, in which a researcher and claimant might interact with the system in a structured dialogue. The system would ask a series of questions about the nature of the claim and the circumstances that favour it. On the basis of this interaction, the claim or set of claims would be clarified. Then the system would suggest some general ways of evaluating the claim, querying the researcher about facilities available and the claimant about the satisfactoriness of the procedure being suggested. After further interaction, a final set of procedures would be recommended. In some cases, the system may be able to provide testing materials itself, including software linking it to a central laboratory facility. Alternatively, the system may have a mode which allows the researcher to input an original procedure, perhaps in consultation with the claimant, which would then be critiqued. The general idea throughout would be to make available to the isolated researcher the collective experience of a set of consultants who could not be consulted directly, for practical reasons. As is the case with other such expert systems, one would have to be careful not to assume its infallibility. All of the caveats mentioned above, and more, would apply, but such a system could serve at least as a valuable initial screening device, especially for researchers in allied areas not that familiar with parapsychological techniques. One potential category of user, for instance, would be mental health professionals who occasionally confront clients claiming psychic powers. To be practical, such a system would need to be made available on one of the public computer subscription services. A side benefit for the parapsychological community of the development of such an expert system is that the expertise itself is likely to become better articulated and organised than is presently the case. Thus even the experts can benefit from the system.

Once such expertise is in place and operating, its existence and reasons for its existence should be communicated to potential research subjects, such that those who do not intend to cheat will regard it as a valuable set of safeguards indicating that the research unit is serious about its work; whereas those who would intend to cheat would be warned away. One word of caution; if the safeguards are described publicly as being absolutely unbeatable, they may well inspire a sense of challenge and source of reputation enhancement for some talented performers. Communicating to potential psychics may involve both general publicity about a research unit and its practices, and specific interactions between researcher and a candidate psychic. Whatever strategy is chosen should be compatible with the unit's philosophy, goals and preferred modes of operation, topics which will be considered below.

2. Pseudopsychics should be persuaded that there could be negative consequences for them if they try and fail. It has been suggested that subjects in parapsychological studies be required to sign a contract, in with both they and the researcher agree to perform certain services and to avoid any deception or misrepresentation to the other. Anyone caught cheating could then be sued for breach of contract. In my opinion such a legally-based approach may have inherent problems, and laws vary considerably among states and countries. Depending on how it was presented, such a contract could be viewed positively, as a realistic safeguard allowing both parties to take comfort in the willingness of the other to sign; or it could increase antagonism and tension by its implication that without such a contract there was insufficient trust. The main issue really is whether any research unit wishes to pursue the possibility of legally based sanctions against a pseudopsychic who can be shown to have been deceptive, and whether the unit is prepared to interact with the legal community to define what laws are relevant and what would constitute sufficient evidence that a law had been broken.

A more easily managed deterrent, one that would not be threatening to any claimant who does not intend to engage in pseudopsi, would be merely to have an open policy that the details of any pseudopsi detected will be published in full, perhaps as a technical note in one of the major journals. No names would be mentioned, but the details of the effect and its means of accomplishment would be revealed. The names of those involved could be made available privately to those with a demonstrable 'need to know'. Thus the research community as a whole is served, and any magician hoping to gain a reputation for cleverness has had the trick exposed, thereby reducing its value.

Perhaps the most important point is that if a research unit wishes to develop a deterrent of any sort and communicate its existence to potential pseudopsychics, that deterrent must be highly selective, e.g. it should deter only pseudopsychics, not the rest of the public who are sincere in their desire to further our collective understanding. For this reason many units may decide to omit deterrents of any sort from their strategy, focussing instead on persuading the pseudopsychic that there just aren't enough benefits to be gained.

3. Pseudopsychics should be persuaded that even if they succeed in deceiving while going undetected, there will be insufficient reward. If one considers the motivations for pseudopsi, one notices that all of them would be aided by extensive publicity for one's psychic achievements, and most of them are quite reliant upon it. How does such publicity get generated? What is the ideal news story from the media standpoint? Based on past media coverage of the field, it appears that strong publicity will accrue to research having the following characteristics: (a) one or two 'superstars' with attractive personalities are involved; (b) the psychic claims being evaluated involve events which are simple, easily understood, dramatic, amenable to graphic representation, and far-reaching in their implications; (c) the research is conducted at a prestigious institution by prestigious researchers; (d) research reports are readily available which describe brief, simple studies with strong results and striking conclusions, based on relatively little research with few people involved; and (e) there is clearly an opportunity for controversy. In other words, research receives media attention if it makes a good story. Pseudopsychics will accomplish many of their goals if they manage to get involved in a good story.

How should we proceed? Any research unit which never generates a good story is, almost by definition, a unit which has failed to learn anything important. Given that we do not wish to forego interesting and straightforward research just to make ourselves less attractive to pseudopsychics, there are still some productive steps we can take to minimize the benefits realisable by a pseudopsychic. One is to maintain a policy of not making public, even in limited distribution reports, the names of any research participants, including not discussing the results of any set of sessions with a specific subject. Such a policy is sensible in providing privacy safeguards for subjects; many will require such privacy anyway, and many others may find that becoming a focus of media attention is more of a harrowing experience than they had expected. The present author once asked a friendly tabloid journalist for guidance on avoiding sensational newspaper coverage; his main advice was, "Don't ever tell me the names of your subjects. Once we know who they are, we can develop stories around them". Since implementation of a privacy policy can occasionally be problematic when confronted by a persistent journalist, the researcher may do well to consult a media or public relations specialist.

A second strategy involves selection of research problems. At one extreme are researchers who specialise in evaluating the unique claims of individual claimants, in hopes of learning about psychic functioning by studying strong performers, especially those who produce striking, easily observed physical effects. Such researchers are at strong risk from pseudopsychics and should make sure that safeguards such as those mentioned above are firmly in place. At the other extreme are researchers who select research problems independently of the claims of individuals and ask questions about the systematic interactions among sets of variables. They are interested in strong effects but are willing to tolerate a certain amount of noise in the data by working with groups of people in order to produce more generalisable effects. Subjects may be presented with a variety of conditions, without always knowing all the characteristics of each condition, in order to learn which condition produces best results apart from subjects' expectations. This style of research, focussing on many subjects and the gradual emergence of a body of knowledge independent of a selected star subject, seems less amenable to exploitation by a pseudopsychic. Perhaps the easiest of all to exploit is the researcher who already 'knows' that psychic functioning exists and is merely trying to find a superpsychic who will produce a concise, publishable demonstration that will impress all who read it. Such researchers will be impressed by isolated, powerful events, and will tend to feel that publication is warranted right away, even though in a sense they have not yet done any real research; they have just found an indication that real research may now be justified. Thus even matters of overall research goals, problem selection and experimental philosophy can involve strategies for minimizing fraudulent activities.

IMPLEMENTATION OF STRATEGIES

Advice on how to implement some of the above strategies can vary so much between research units that it is fair to say that each unit should bear the responsibility for formulating its own general policy and specific strategies, with the help of available sources of information, consultants, and so on. Some research philosophies provide more opportunities for pseudopsychics to fulfil their needs than others. In general, pseudopsychics thrive on bad science, at both the conceptual and procedural levels.

There are two additional aspects of strategy implementation favoured by the present author. One is that if approached by a claimant, one can explore with the claimant how their apparent skills may match a research question under active exploration by one's own unit, such as to plan how the claimant may be able to practice slight modifications of these skills even on their own, such that they can match and (ideally) succeed at one or more of the procedures being employed with confidence at one's own unit. This can be done with some flexibility, i.e. it need not involve a forcing of an individual to adopt a sterile, uninteresting procedure. In fact, a major challenge for researcher and attendant experts is to develop imaginative yet sound procedures for measuring psychic functioning, procedures that will accommodate the interests of a variety of serious claimants. If this strategy is followed, ideally the pseudopsychics will drop out once they realise what they are confronted with, whereas serious claimants will persist on their own such that when they actually participate in controlled laboratory research they are ready for it.

A second aspect of strategy implementation favoured by the present author is the use of networked computer systems to allow claimants with access to microcomputers in their own homes, workplaces, leisure spots and so on, to interact experimentally with a central parapsychology laboratory on-line, using a variety of computer-assisted protocols, whenever and under whatever personal conditions they wanted, without a loss in experimental safeguards. A prototype of such a system has already been developed and more will likely follow as costs decrease. By using one of these systems, one can address the needs of the serious claimant who is put off by the necessity to work in a laboratory environment, while at the same time weeding out the pseudopsychics who can no longer use complaints about laboratory sterility as an excuse for seeking weaker conditions.

In any attempts to minimise the problem of subject fraud in controlled research, it is hoped that researchers will take the precautions needed to weed out pseudopsychics without becoming overly concerned or obsessed about their presence. In principle they can never be completely eliminated, but in practice there is much we can do to ensure that our research is productive and enjoyable for all concerned, while a particular form of unproductive pest simply dies out from lack of attention and nourishment.

ABSTRACT

In recent years, parapsychologists have become increasingly concerned with the detection and elimination of fraud committed by pseudopsychics, both in field work and controlled laboratory investigations. The present paper focusses on the need to understand the motivations of pseudopsychics, so that research units can develop strategies for minimising the likelihood that pseudopsychics will have any interest in them. It is argued that researchers should acquire expertise in the general strategies of pseudopsi and in the specific tactics that are germane to whatever claims are currently under investigation, should avail themselves of guidelines and practices developed at other laboratories, should consult professional magicians and other experts where appropriate and may even wish to interact with computer-available expert systems developed by interactions among knowledge engineers and relevant experts. Such expertise should be involved in the research protocols of the laboratory, and potential claimants should be made aware of the existence of such expertise. Individuals or groups may also wish to develop policies that will serve as deterrents for pseudopsychics, such as a policy of full disclosure of the details of any detected fraud. Researchers are also encouraged to consider whether their general patterns of research are unusually conducive to the sort of media attention that suits the needs of the aspiring pseudopsychic. Some suggestions for specific

strategies for discouraging pseudopsychics while still encouraging others is given, and a short reading list is provided.

SUGGESTED READINGS

Cummins, J.J. (Ed.) 'Annemann's practical mental effects', New York, Louis Tannen's Magic Shop, 1944.

Christopher, M. 'Mediums, mystics and the occult', New York, Thomas Y. Crowell, 1975.

Corinda, T. 'Thirteen steps to mentalism', New York, Tannen Magic, Inc., 1968.

Fitzkee, D. 'The trick brain', Oakland, Calif., Magic Limited, 1944.

Fitzkee, D. 'Magic by misdirection', Oakland, Calif., Magic Limited, 1945.

Fuller, U. (Gardner, M.) 'Confessions of a psychic: The secret notebooks of Uriah Fuller', Teaneck, N.J., Karl Fulves, 1975.

Fuller, U. (Gardner, M.) 'Further confessions of a psychic', Teaneck, N.J., Karl Fulves, 1980.

Hyman, R. 'Cold reading: How to convince strangers that you know all about them', The Zetetic, 1, Spring/Summer 1977, 18-37.

Kaye, M. 'The handbook of mental magic', New York, Stein and Day, 1975.

Keene, M. 'The psychic mafia', New York, Dell, 1976.

Kurtz, P. (Ed.) 'A skeptic's handbook of parapsychology', Buffalo, N.Y., Prometheus Press, 1985, 177-356.

Mills, J. 'Six years with God: Life inside Rev. Jim Jones's peoples temple', New York, A & W Publishers, 1979.

Morris, R. 'What psi is not: The necessity for experiments', in Edge, H., Morris, R., Palmer, J., and Rush, J. 'Foundations of

Parapsychology', London: Routledge and Kegan Paul, 1986, 70-110.

Triplett, N. 'The psychology of conjuring deceptions', American Journal of Psychology, 11, 1900, 439-510.

R.L. Morris,
University of Edinburgh,
Department of Psychology,
7 George Square,
Edinburgh, EH8 9JZ
U.K.

EUR.J.PARAPSYCHOL., V.6, 1986

A PHYSICAL THEORY FOR PARANORMAL PHENOMENA

Giroldini William

The existence of a physical paradox named 'nonlocality' or 'nonseparability' has been known for a long time. The paradox was first drawn to the attention of the scientific world by A. Einstein, B. Podolsky and N. Rosen (1935), and from that time, it has been the centre of increasing interest on account of its profound scientific and philosophical implications (Bohm and Hiley, 1975; d'Espagnat, 1979). Attention to this problem has only occurred in recent years, following the carrying out of numerous specially designed experiments which have confirmed the paradox's existence and the quantum predictions (Aspect, Dalibard and Roger, 1982). Since nonlocality allows interesting speculations about the nature of paranormal phenomena, it is necessary to describe briefly the meaning of this paradox.

Nonlocality is implied, for example, also in the wave functions which describe the behaviour of the electrons around the atoms and molecules, yet it takes physical importance in particular when it is described as a direct consequence of a spin conservation law. Spin is a fundamental property of elementary particles, and can be described approximately as a rotatory motion of elementary particles around themselves. The spin is quantified, i.e. it can only assume multiple integer values of the quantity $h/4\pi$, where h is Planck's constant. For example, the electron and the proton have spins equal to $h/4\pi$, while the photon has spin $h/2\pi$ and the neutral pion has zero spin. It has been found experimentally that not only the spin vector is quantified, but also its projection (or component) along a definite direction of space. For example, in the most simple case, a particle with spin $h/2\pi$ can be oriented parallel, antiparallel or transversal with respect to the direction of motion.

Now, elementary physical processes exist in which couples of particles having correlated spins can be created. For example, when a neutral pion spontaneously decays, it can emit two photons going in opposite directions and with antiparallel spin, that is to say, if one photon has spin $+h/2\pi$, the other has spin $-h/2\pi$. This fact derives from the conservation law of total spin of any elementary physical process, since, as stated, the spin of the neutral starting pion was zero. If we, by means of a suitable process, change the direction of the vector spin of one of the created photons, the conservation law of total quantified spin also requires that the spin of the other photon undergoes a variation of direction, which exactly compensates that of the first photon. Now, the crux of this paradox is that this 'answer' from the other particle at distance to the action exerted on the first particle, must manifest itself instantaneously and independently of the distance which separates the two particles and of the elapsed time from their own creation. For those interested, we can say that it is possible to change the direction of a photon spin by means of a simple polarizer filtre. When a photon passes through a polarizer filtre, its spin alligns itself parallel to that of the filtre, independently of its initial spin orientation.

Since the 1970's, numerous experiments have been performed which confirmed the predictions of quantum theory, from which the paradox was born, including experiments which have eliminated the possibility of signals being transmitted between two correlated particles at speeds less than or equal to the speed of light. Thus, quantum formalism shows that nonlocal correlations exist, but also that they are independent of whether or not the microphysical correlated systems are observed. The creation, by different means, of couples of correlated particles does not concern only the photons, but also includes, for example, electrons, and this is in every respect the most interesting case. In fact, the peculiar properties of nonlocality, that is the interaction at a distance without any attenuation, the instantaneity and the specificity of the bond between the two correlated particles may constitute, in our opinion, the original physical basis of phenomena which we call 'telepathy' and 'psychokinesis'.

In recent literature, many authors have proposed the idea that quantum nonlocality may serve as a model for paranormal phenomena. We can quote E.H. Walker (1973, 1975), W.G. Roll (1979), C.N. Villars (1981, 1983) and finally J.E. Charon (1983). At the end of this paper, some differences between Walker's model and our model will be discussed briefly. Let us see, then, the means by which the transition from a series of microscopical phenomena to a macroscopical paranormal phenomenon could be realized.

THE THEORY

We now assume that a 'bosonic couple' (b.c.) is a couple of electrons with total spin zero or equal to $h/2\pi$. At the moment of bosonic couple creation, the two particles are obliged to be very near both in space and time. Now we introduce the following two postulates which constitute the basis of our theoretical model.

A The modulus of the total spin vector of the bosonic couple never changes; the change in the vector spin of one member of the couple implies that the spin of the other particle must change instantaneously in order to keep the total spin unchanged.

B The total spin vector possesses $2S_T + 1$ possible directions with respect to the direction which is defined at each moment by the total impulse vector: $\vec{P}_T = \vec{P1} + \vec{P2}$

If, for example, the bosonic couple is constituted of two electrons, their total spin $h/2\pi$ can be oriented parallelly or antiparallelly or perpendicularly with respect to the total impulse vector.

Let us examine the consequences of the two postulates A and B. For simplicity, we suppose that a couple of correlated electrons (b.c.) have been generated by a suitable process, and an elastic collision between one of the two electrons of b.c. and a third free electron or atom takes place. In a description of the event without postulates A and B, the interaction cannot, in any way, influence the trajectory and the energy of the third electron which is an arbitrary distance from the first two, but in view of these postulates this conclusion is not true any more. As a consequence of the collision, the direction of the vector spin of the b.c. and the spin of the colliding electron may change and analogously may change the impulse vectors of all the particles involved. The possibility of spin vector variation of the b.c. in the collision implies that all three particles, and not only the two directly involved, can undergo changes of impulse and kinetic energy as a direct consequence of postulates A and B, although the total impulse, energy, and spin are fully conserved. The interaction does not destroy the bosonic couple, but may exchange energy and impulse among their components. The exchange is virtually instantaneous in time and independent of the distance from the two correlated particles.

In order to carry on the model explanation, it is now necessary to describe the physical mechanisms suited to create electronic bosonic couples in normal biophysic conditions of living systems, and also the mechanisms which bring about their destruction, without the violation of postulates A and B. A possible hypothetical mechanism can be linked to the almost free electrons, which are present in nearly all materials. They originate from intrinsic properties of materials, or from reticular flaws, impurities, etc. and substantially they are responsible for the electrical conductivity of the same materials. If three free electrons collide elastically, a small but not zero probability exists that, at the end of the interaction, a couple of correlated electrons are created while the third electron remains free. Ihis is only one of the possible ways for b.c. creation. Vice versa, a mechanism of b.c. destruction can be described by means of a b.c. one member interaction with another free electron; sometimes, at the end of the interaction, a photon can be emitted which continues the spin of the b.c.. In the final state, we have three electrons with uncorrelated spin, and one photon having spin $h/2\pi$. All these hypothetical mechanisms are substantially in accordance with the conservation laws.

After the description of the possible mechanisms of generation, propagation, and destruction of bosonic couples, it is necessary to see in which way we can arrive at the macroscopic phenomena of telepathy and psychokinesis. It is possible to imagine different mechanisms, but it is important to show that the proposed mechanisms are not at variance with current biophysical knowledge and that they are experimentally verifiable.

TELEPATHY

We propose, as a general idea, the following mechanism for telepathic phenomena between two persons. In the time preceding an ESP event, the two people have had physical contact so that a lot of electronic bosonic couples were separated and one of the two b.c. particles was transferred to the other person. Afterwards these particles (electrons) were 'stored' or 'kept' in cerebral areas of memory, together with the biochemical mechanisms of recording experiences in the brain. When one of the people evokes, either consciously or unconsciously, episodes, images etc. which are stored in that area of memory, he can cause a physical action on the separate particles of the b.c. which are in that area. At the same time, due to postulates (A) and (B), the other members of separate b.c.'s, always present in analogous cerebral structures of the other person, react and cause a chain of biochemical events, which can culminate in the activation of those memory areas. In extreme cases, the generation of real and conscious hallucinations are possible. This mechanism might seem very complex and artful and for this reason we think it would be useful to examine carefully the various points of the model.

a) The separation of b.c. components, due to the physical contact of a person with either objects or other people, is very reasonable, because the two b.c. electrons are not obliged to remain near each other: their bond is independent of distance (postulate A).

b) At present the molecular processes of memorisation of experiences caused by events in the external world, as well as the process of information recovery, are unknown. Nevertheless we think that at least some mechanism must exist which transforms the information into reversible (or irreversible) modifications of molecules which are present in the nervous cells. If, during the memorisation step, an electron of a separate b.c. arrives at this specific nervous site, it can be 'captured' by molecules involved in the memorisation process. The captured electron must be confined in those particular molecules for an indeterminate time and with a very low probability of interaction with other free electrons.

c) When the molecular information is 'read' (or recovered), this electron must participate in some biochemical process (e.g. oxo-reduction) which contemporaneously transfers energy to another remote electron of a b.c., wherever it is. In biological systems there exist many molecules which are potentially able to capture free electrons, stabilizing the captured electrons in suitable atomic or molecular orbitals. For example: quinones, N-alkylpyridines, disulphides bridges, enzymes containing transition metals, and so on. In the first step the captured electron gives a free-radical; in the next step the electron becomes part of a chemical bond between two atoms: C-H, C-O, C-N etc. In this condition, the electron is indefinitely stable at least until the molecule is 'read' (as considered before) or destroyed. It is possible to argue that, after its birth, the b.c. can interact with billions of other particles within a few seconds, and therefore the b.c. could lose its identity very rapidly. The loss of identity (i.e. loss of information) of the b.c., though, cannot happen by simple elastic collision with other electrons, atoms or molecules; in fact the conservation of identity of the b.c. is assured by the laws of spin conservation (postulate A). The billions of collisions within a few seconds with other particles, can only vary the direction of the bosonic spin vector and distribute the kinetic energy of the involved particles again.

The only process which brings about the destruction of a b.c. is that which foresees the interaction with another free electron, and the emission of a photon at the end of the interaction. The probability of this type of event depends upon the concentration of free electrons in normal living matter, on the kinetic energy of involved electrons, and on the intrinsic features of the process. The number of free electrons in living matter is normally many orders of size lower than the number of electrons engaged in chemical bonds in molecules or ions. Moreover, if the event in question (collision with emission of a photon) needs a minimum threshold of energy for it to happen, being comparable with the energy (activation energy) request for many chemical reactions ($\sim$10 Kcal/mole) then the probability of this event, at a temperature of 37°C, can be so little as to allow the survival of a b.c. for many hours. This time is sufficient that a part of these electrons may finally be captured by molecules involved in the process of memory. The electron of a b.c. which is part of a covalent chemical bond, is protected from the loss of information owing to the high quantity of energy necessary to break a chemical bond (typically 40-100 Kcal/mole). In the biological system, in fact, the formation and the breaking of chemical bonds, is generally controlled by highly specific enzymes which act in well-defined periods of time.

e) A very important point is the description of possible ways by means of which the free electrons belonging to a certain object or person can be captured in specific memory areas of another person, with sufficient selectivity with respect to the free electrons coming simultaneously from a different source. It is possible to think that the 'selectivity factor' is proportional to the number of physical contacts of a person with a certain target and with the degree of emotion or interest that the target stirs up in the person. We know, of course, that the greater part of daily events are regularly forgotten after a short time, and only the more emotional situations leave lasting memories and sensations. The consequence is that the entrapment processes of b.c.'s are in dynamic equilibrium with b.c. removal processes due to the normal metabolic exchange of molecules and enzymes in the biological system. Therefore, only frequent and important contacts between two persons (e.g. mother/son, husband/wife etc.) should bring a sufficent b.c. concentration to produce an ESP event. This model foresees that sporadic, non repeated contacts with several and different persons, have a low probability to produce an ESP event (Persinger, 1974).

f) In a typical situation, if a son is in great danger of losing his life, and at this moment thinking intensely about his mother, then the mechanism of nonlocal connections could cause the activation of corresponding cerebral areas in his mother, who could feel a sensation of uneasiness, anxiety or actual physical danger about her son. If the cerebral stimulation is considerable, she could have a hallucination in which she 'sees' the son in danger. It is known that a mechanical or electrical stimulation of small cerebral areas of memory (few square millimeters) can produce intense and whole hallucinations, as in a film, in which the subjects lives again episodes and images forgotten for years (Penfield, 1975).

g) It is possible to calculate that the single quantity of energy transmitted or received with the proposed mechanism is about 10^{-11} erg, that is the average energy of a typical chemical bond in organic molecules. This quantity must be multiplied by the number of particles involved. It follows that the total energy involved can sometimes increase the excitation threshold of a neuron network. If this threshold is exceeded, it can trigger complex associative mechanisms which eventually brings the perception to a conscious level. If the threshold is not exceeded though, unconscious effects can be produced, in particular psychosomatic variations of humour, heart rhythm, electrical skin resistance, etc.. In short, telepathy could be seen as a partial reproduction in a brain of the excitation patterns of another brain. A corollary of this conclusion is that in omozygote twins, we should observe telepathic events in a significantly higher number than compared to average couples of people, on account of the great structural similarity of the two nervous networks involved.

PSYCHOKINESIS

Psychokinesis originates as a natural consequence of postulates (A) and (B). For estimating the forces and energies in a typical example of psychokinesis, let us suppose that the PK event concerns an object having a mass of about 10 grams, and that it moves with a final velocity of 20 cm/sec. We suppose also that its motion continues for inertia and stops because of attrition within about one second. On the hypothesis that the kinetic energy acquired by the object is derived from the kinetic energy of a b.c. a simple estimate brings us to the conclusion that the contribution of one electron from about every 10^9 - 10^{10} electrons is necessary to justify the energy balance of the event.

Moreover, we observe that, in inanimate objects, the b.c. storage process can be different to that of living systems. In particular, metals in the first and second series of transition, which have suitable external orbitals, could be implied for receiving a single electron in atomic half-empty orbitals. These elements (e.g. Cr, Mn, Fe, Ni, etc.) in suitable concentrations, that is less than one part per million, are present in any material both living or not and could be selective traps for free electrons, including the bosonic couples.

A possible outline for PK phenomena could be the following: When b.c. particles are put into action in the brain of an agent, in connection with his thoughts turning to the object, the corresponding particles present in the latter transfer energy and impulse to the same object. If the resultant of all the single impulses is sufficiently different from zero, then a force of displacement of the object in a direction of space derives. Normally, given a random distribution in space of a large number of impulse vectors, the resultant is statistically near zero. An important circumstance exists though which can yield a non-zero result. The cerebral neurons and their prolongations (axones and dendrites) are organized according to the high spatial order, not casual, and inside each nervous cell many enzymes are permanently fixed, for example on the cellular membrane with stable and definite orientation with respect to it. Therefore, when the enzyme creates or breaks chemical bonds, the impulse vectors associated with the electronic transitions can have a non-random resultant which is defined in a particular nervous network. It is evident that this mechanism requires the simultaneous presence of many conditions, each of which is fairly improbable: a sufficient concentration of separate b.c.'s in the body of the agent and the object, a short time since their activation in the agent, and a resultant of the impulses not zero in the object.

The proposed model can easily explain that singular and rare phenomenon which is the self-combustion of inflammable objects in the presence of a poltergeist. In this case, we can argue that the activation of electronic b.c.'s on the surface of an object, owing to the unconscious processes of activation in the brain of the agent, can create so-called 'free-carbon radicals'. The free radicals are chemical molecules having very high reactivity, which by reacting with oxygen in the air develop heat and can start self-combustion.

The model based on non-locality seems to explain very well quite a lot of paranormal phenomena and also their extreme rarity, which can be explained by the peculiar conditions required for their expression. The classes of phenomena best explained by this model are the following:

1) Telepathy, dreams and telepathic hallucinations between acquainted persons, apparitions.
2) Clairvoyance with possible telepathic components.
3) Psychometry with possible telepathic components.
4) Sudden movement of small objects, noises, or 'raps' for example in poltergeist.
5) Psychical action towards other persons, animals, plants and electronic circuits, micro PK of 'metal benders' (Hasted, 1980).
6) Self-combustion of inflammable objects in presence of poltergeist.
7) 'Psychical impregnation' of objects and places (Ryzl, 1965, 1967). This theory supplies a precise and definite mechanism for this strange and debated phenomenon, also observed in the laboratory.

On the contrary, this theory does not seem to fit the following cases:

8) Psychokinesis of massive objects, levitations.

9) Precognition, pure clairvoyance, pure psychometry.
10) Processes of materialization/dematerialization of objects.
However, it is important to observe that phenomena (8) to (10) were never observed under laboratory conditions. Another observation is that the non-local connection between the two electrons could imply a kind of precognition, owing to the time symmetry of the equations describing the two correlated particles. Nevertheless the problem is complex and it will not be further discussed in this paper.

THE MODEL PREDICTIONS

The experimental control of theory is based, above all, on physical experiments which must verify the existence of the foreseen 'bosonic couples' in accordance with postulates (A) and (B). The project and the realization of these physical experiments is beyond the aim of this paper. Nevertheless, even after a possible confirmation of the two postulates, it is plain that considerable difficulties exist for verifying experimentally the further hypothetical mechanisms for ESP/PK events. Let us examine some specific predictions, which are deducible from the proposed model:

a) The success rate in standard telepathy tests ought to be higher if the subjects have met before the test and have gone through important experiences together.
b) The success rate in clairvoyance tests using Zener cards, psychometry of objects, psychokinesis with dice or electronic RNG generators, ought to be higher in the case the subject has touched the targets before the tests.
c) According to the proposed model, ESP/PK is independent of distance.
d) If the ESP/PK experiment goes on in time without renewed physical contact of the target with the subject then a progressive diminuition of the success rate should be observed until a purely casual level is achieved.

The predictions a, b, and c are equal to those deducible from the model proposed by C.N. Villars (1983). On the other hand, the prediction d is original and derived from the probable destruction of bosonic couples when put into action in cerebral areas of the agent: acting repeatedly on the same target the number of surviving b.c.'s must eventually become too small to yield significant ESP/PK results. From the experimental point of view, prediction d seems very difficult to distinguish from a parallel, principally psychological cause based on weariness and diminuition of enthusiasm due to the continual repetition of the same experiments.

Other conclusions from the proposed theory are:
a) The brain, in which this kind of at-a-distance correlation shows itself, already contains all the information sufficient for elaborating generally unconsciously the 'message': the ESP event consists only of a selective excitation of those cerebral areas (Servadio, 1967; Krippner, 1967; Hernandez-Peon, 1967).
b) There is no transmission-reception of some new information as in radio communications.
This model allows us to overcome the difficulties associated with an information theory approach, for example the problems of codification, transmission, reception, and decodification of the information.

As previously described, we assume that the connection between an object or a person and a 'psi-agent' is realized by the storage of b.c. electrons in certain cortical areas (memory). Walker's model, instead, introduces such a nonlocal connection by a feedback loop of synaptic processes in the brain, in concomitance with external physical processes, by nonlocal 'hidden variables'. Walker's theory is fundamentally an 'observational theory'; he hypothizes, in fact, that the conscious act of observation of a given event can influence, at least partially, the observed event by the hidden variables. If these hidden variables are under the partial control of the consciousness and the will, then there is the possibility of influencing the wave vectors which describe the state of a physical system. Our model does not take into consideration the hidden variables and assumes that the nonlocal connections exist independently of whether or not their effects are observed. However, there is no fundamental incompatability between the two theories and we think that they can describe two different and comparatively independent ways of mental action upon physical reality.

It is possible to conceive an experiment where the two models lead to different predictions. The experiment is based on two identical random event generators utilizing the noise source of a Zener diode, as described by Nelson, Jahn and Dunne (1986). One generator must be completely isolated electrically from the ground, whereas the other has the ground terminal connected by a metalic wire to a subject who must try to influence mentally both generators in a same selected direction (PK+ or PK-). Each generator provides some sort of feedback, usually a visual display, that tracks the degree of shift from the baseline distribution. According to the observational theories the success rate of psychokinesis on the two random number generators should be substantially identical, whereas according to our model a more significant result should be obtained on the generator electrically connected to the subject since only in this condition an exchange of b.c. electrons is possible between the person and the object. Naturally the subject should not know that there is a difference between the two generators.

Finally we propose an experiment which should allow the experimental verification of both postulates A and B, independently of the mental action of a person (which is always difficult to reproduce). We suppose to have a piece of conductor material, such a graphite, metal, silicon, etc.. This piece is kept in a state of high electrical insulation from surrounding places for at least 48 hours. Then the piece is divided in two equal parts one of which is connected with a thermocouple which measures continually the temperature (the thermocouple is electrically insulated and the piece is kept at the temperature T1 with appropriate thermic insulation). The second piece of material is then heated to the temperature T2, with T2>>T1. If the two postulates A and B are true then a spontaneous increment in temperature in the first, unheated, piece should be observed. The degree of temperature increment is about:

$$\triangle T1 = \triangle T2 \cdot C_{p2} \cdot R \cdot J \,/\, C_{p1}$$

where:

$\triangle T1$ = increment in temperature in the first piece of material.
$\triangle T2$ = increment in temperature in the second piece of material.
C_{p1}, C_{p2} = specific heat of material at the temperatures T1 and T2.
R = ratio between the number of b.c. electrons and the total number of electrons in the material. We think that R is in the range of 10^{-5} - 10^{-10}.
J = yield of energy transfer of process. We value that J is about 0.1.

Since the specific heat of all the substances decreases when the degree of temperature lowers, it is best to cool the first piece of material to 4.2 $^{\circ}$K in order to obtain a value of $\triangle T1$ experimentally measurable. A suitable temperature for T2 is about 1300 - 1500 $^{\circ}$K, which must be achieved within one or two seconds; for example with graphite or a metallic powder (Zn, Fe, Mn etc.) it is possible to realize a very fast combustion, mixing these substances with potassium perchlorate ($KClO_4$) and igniting. This experiment is crucial for the theory: the major uncertainties come from the calculations of the values of R and J, based on complicated considerations of physics.

ABSTRACT

A physical theory for paranormal phenomena is proposed, in particular for telepathy and for psychokinesis on small entities. The starting point of the model is the phenomenon of 'non-locality', which in recent years has been confirmed by numerous specifically designed experiments. Briefly, non-locality implies the existence of a new kind of at-distance interaction between two particles based on the spin conservation law.

Two physical postulates are introduced by means of which the author describes the transition from a series of microscopical physical phenomena to a macroscopical paranormal phenomenon. The model proposes the hypothetical existence of electron couples having a well-defined spin correlation (bosonic couples). The author proposes a general mechanism for telepathy based on the idea that in the period preceding an ESP event two people have had physical contact so that a lot of electronic bosonic couples were separated and one of the two bosonic couple particles was transferred to the other person. Afterwards these electrons are 'stored' in cerebral areas of memory together with the biochemical mechanisms of recording experiences in the brain. When one of the persons evokes, either consciously or unconsciously, episodes, images, etc., which are stored in that area of memory this can cause a physical action on the separate particles of the bosonic couples which are in that area. At the same time, according to the two postulates, the other members of the separate bosonic couples always present in analogous cerebral structures of the other person react and cause a chain of biochemical events which can culminate in the activation of those memory areas. In extreme cases, the generation of real and conscious hallucinations is possible.

In short, telepathy could be seen as a partial reproduction in a brain of the excitation patterns of another brain. Such a mechanism could also explain psychokinesis of small entities and, in general, the direct influence of mind upon matter, living or non-living. Some testable predictions are deduced from the proposed model.

ACKNOWLEDGMENTS

The author is grateful to Dr. R. Rosa, Prof. F. Bersani (University of Bologna) and Dr. M. Biondi for their helpful criticisms and suggestions.

REFERENCES

Aspect A., Dalibard J., Roger G. 'Experimental test of Bell's inequalities using variable analysers', Phys. Rev. Lett., 1982b, 49, 1804-1807.

Bohm D., Hiley B. 'On the intuitive understanding of nonlocality as implied by quantum theory', Foundations of Physics, 1975, 5, 93-109.

Charon J.E. 'L'esprit et la relativite complexe', Editions Albin Michel, 1983, p 211.

d'Espagnat B. 'Quantum theory and reality', Scientific American, 1979, 241, 128-140.

Hasted J.B., Robertson D. 'Paranormal action on metal and its surroundings', J.S.P.R., 1980, 50, 784.

Hernandez-Peon R. 'Stati di coscienza alterati: un modello unitario', in Cavanna R. 'Aspetti scientifici della parapsicologia', Editore Boringhieri, 1978, p 191.

Krippner S. in Cavanna R. 'Aspetti scientifici della parapsicologia', Editore Boringhieri, 1978, p 115-116.

Nelson R.D., Jahn R.G. and Dunne B.J. 'Operator-related anomalies in physical systems and information processes', J.S.P.R., 1986, 53, 261-285.

Penfield W. 'The mystery of the mind', Princeton University Press, 1975.

Persinger A.M. 'The paranormal', Laurentian University, Sudburry-Canada; Mss Information Corporation 655 Madison Avenue N.Y. 10021, 1974, p 106-142.

Roll W.G. 'Psi structures', Theta, 1979, no 2.

Ryzl M., Barendregt J.K., Barkema P.R., Kappers J. 'An ESP experiment in Prague', J.o.P, 1965, 29.

Ryzl M., Otani S. 'An experiment in duplicate calling with Stepanek', J.o.P., 1967, 31.

Servadio E. in Canvanna R. 'Aspetti scientifici della parapsicologia' Editore Boringhieri, 1978, p 117.

Villars C.N. 'Quantum nonlocality', Psychoenergetics, 1981, 4, p 15-23.

Walker E.H. 'Foundations of paraphysical and parapsychological phenomena' in Oteri L. (ed) 'Quantum Phy. and Parapsych.' , N.Y., 1985.

Walker E.H. 'Consciousness in Quantum Theory', J. Study Consciousness, 1972-1973, 5, 2, 257.

Dr. Giroldini William,
via Pace 6,
20097 S. Donato,
Milano,
Italy

EUR.J.PARAPSYCHOL., V.6, 1986

PARAPSYCHOLOGY AND THE SENSORY-MOTOR HYPOTHESIS

Magne Dybvig
University of Trondheim

PART ONE
ON THE PHILOSOPHICAL IMPORTANCE OF PARAPSYCHOLOGY.

PARAPSYCHOLOGY AND PHILOSOPHY

There is a widespread popular interest in what may vaguely be described as inexplicable, mystical or supernatural phenomena. Such occult phenomena, as I shall call them, can also be made the object of scientific research. One can then distinguish between two different approaches (which can be combined in various ways). On the one hand, seemingly occult phenomena may be studied in order to find out why people believe in such phenomena. This is the historical, psychological, antropological and sociological study of superstition, witchcraft, magic, and so forth.

On the other hand, occult phenomena may also be studied in order to find out whether or not beliefs in such phenomena contain some truth -that there really are (something more or less similar to certain types of) occult phenomena. This is what we may call the validity-oriented study of occult phenomena.

No doubt parapsychology constitutes the most interesting example of a validity-oriented study of occult phenomena. By 'parapsychology' I understand the investigation of psychological phenomena with a view to

1) ascertain whether or not some of these phenomena are paranormal in the sense that they involve a so-called psi-component, i.e., information about, or influence on, the environment that does not make use of any known sensory or motor mechanisms, and
2) formulate hypotheses about the nature of this psi-component and how it is related to other phenomena. Parapsychological research is validity-oriented research in so far as it involves investigating if at least some (seemingly) occult phenomena -such as telepathy, prophetic dreams, out-of-the-body-experiences- are genuine phenomena in the sense that they contain a psi-component that cannot be explained by science at the present time.

Parapsychology has remained a small academic discipline which has many of the characteristics of a rejected science, living a life on the borderline between what for most scientists are clear cases of pseudosciences, like astrology and numerology, and what are clear cases of respectable, recognised sciences (note 1). The total number of researchers engaged in parapsychological research is small, little of this research is published in 'orthodox' scientific journals, public financial support is negligible, and so on (cf. Allison 1979). Nevertheless, parapsychology has developed many of the typical traits of a scientific speciality. There are journals that specialise in publishing parapsychological research, parapsychological institutes and laboratories inside and outside recognised academic institutions, societies supporting parapsychological research, research- and teaching appointments outside and inside recognised academic institutions, and a professional organisation (PA) for parapsychologists with professional qualifications (cf. Collins & Pinch 1979).

Parapsychology has also attracted the attention of philosophers. This has partly been a matter of direct engagement in parapsychological research and in institutions supporting such research, partly a matter of philosophical research connected with parapsychological research. Throughout the years the professional philosophical interest in parapsychology has been quite considerable, considering the (comparatively) small numbers of philosophers. Several internationally well-known philosophers have concerned themselves with parapsychology (note 2).

Most philosophers with a sympathetic interest in parapsychology could be said to agree on the following points, I believe:

Parapsychological research has at least

a) succeeded in casting serious doubt on the proposition that paranormal phenomena never occur, and has consequently

b) succeeded in casting serious doubts on scientific or philosophical theories that imply, or presuppose, that such phenomena never occur, and has consequently also

c) succeeded in making it philosophically and scientifically interesting to contemplate the possibilities of theories that are compatible with, or even imply, the occurrence of such phenomena.

Some philosophers with an interest in parapsychology have maintained that parapsychology has done more than cast serious doubt on the proposition that paranormal phenomena never occur, it has also made probable, or even proved, that such phenomena really occur. Personally I am not willing to go further than the minimum expressed in a - c, but for me that is sufficient for taking an interest in parapsychology as a validity-oriented scientific study of occult phenomena.

SOME CENTRAL PARAPSYCHOLOGICAL CONCEPTS AND HYPOTHESES

The central parapsychological phenomena are best understood as various forms of exceptions to certain generally accepted hypotheses about how human beings (and other organisms) interact with the environment. Consider the following assumption, which we may call the sensory hypothesis, SH:

Information about the environment presupposes, directly or indirectly, sensory information about the environment.

Any exception to SH implies the combination: Information, but not (direct or indirect) sensory information. This combination means that we have a paranormal phenomenon, namely extrasensory perception, ESP. The sensory hypothesis is part of a larger theory of how human beings (and other organisms) adapt to the environment. The main elements of this theory may be formulated as follows.

1. Experiences and actions may in various ways, and to various degrees, correspond to, or fit, what goes on in the environment. A thought may be true or false and in that sense correspond to what is the case; what one wishes would happen, may happen and in that sense correspond to what actually happens; a depicted scene may be more or less similar to a real scene and in that sense correspond to a real scene, and so forth. That experiences and actions may thus correspond to what is the case, seems to be an important part of our common sense and scientific picture of ourselves (and other organisms). It is part of depicting ourselves as situated in an environment that we can either influence (and so make the environment conform to, or fit, what one wishes or intends) or obtain information from (and so make one's thoughts, images, assertions etc., conform to the environment).

2. Within the class of such organism-environment correspondences (oe-correspondences, for short) it is possible to distinguish between those correspondences which are due to chance and those correspondences which are not due to chance. It may for instance be just a coincidence that someone's wishes correspond to what actually happens, but if the person in question is capable of goal-directed actions, then such correspondences will not usually be the result of pure luck. In the same way, if a person is capable of perceiving what goes on in the environment, then most of the time it will not be just a coincidence that there is a correspondence between what the person experiences and what actually goes on in the environment. Generally, a precondition for looking on ourselves as conscious beings capable of intentional actions, is that we are thus capable of distinguishing between oe-correspondences which are, and oe-correspondences which are not, a result of coincidences.

3. Oe-correspondences which are not accidental corerspondences either reflect the fact that we have information about the environment and/or that one influences the environment so that a correspondence is the result.

4. In so far as a oe-correspondence is the result of information about the environment, it has come about in certain characteristic ways:

a) It is a result of one now perceiving something, or
b) it is a result of one having perceived something and still remembering it, or
c) it is a result of present or earlier perceptions which generate expectations and/or associations which somehow correspond to what is,

or what has been, perceived, or
d) it is a result of inferring what is the case on the basis of what one perceives, or has perceived, or expects to perceive, or associates with things already perceived.

5. In so far as an oe-correspondence is the result of influencing the environment, then it is either:
a) the result of bodily activities which directly bring about changes in the environment, or
b) the result of bodily activities which more or less indirectly bring about changes in the environment.

Let us call 4a - 4c sensory information (Se). Sensory information then includes both direct and indirect sensory information. 4 expresses what I call the sensory hypothesis (SH): All cases of information about the environment are cases of (direct or indirect) sensory information about the environment.

Let us call 5a - 5b motor influence (Mo). 5 then expresses what I shall call the motor hypothesis (MH): All cases of influencing the environment are cases of (direct or indirect) motor influence on the environment.

The general hypothesis that non-accidental oe-correspondences are the result of sensory information and/or motor influence, I shall call the sensory-motor hypothesis (SMH): All cases of non-accidental oe-correspondences between subject and environment are either cases of sensory information about the environment and/or cases of motor influence on the environment.

As mentioned earlier, central parapsychological phenomena may be regarded as exceptions to general hypotheses like SH, MH and SMH. Exceptions to SH I shall call extrasensory perception, ESP. The hypothesis that such exceptions really occur, I shall call the ESP-hypothesis (ESPH). I take it that an important task for parapsychological research is to decide whether or not ESPH is in fact valid. Since we have, as trivial logical equivalences, both
1) SH if, and only if, not-ESPH, and
2) ESPH if, and only if, not-SH
this means that one of the important tasks of parapsychological research is to decide whether or not SH is true. To decide between ESPH and not-ESPH is the very same thing as deciding between not-SH and SH, just as deciding between SH and not-SH is the same thing as

deciding between not-ESPH and ESPH.

Exceptions to the motor hypothesis (MH) I shall call psychokinetic phenomena, PK. The hypothesis that such phenomena occur I shall call the PK-hypothesis (PKH). Again I take it that it is an important task for parapsychology to decide whether PKH is true or whether not-PKH is true. That is, since we have the logical equivalences
1) MH if, and only if, not-PKH, and
2) PKH if, and only if, not-MH
one can say that one of the central tasks confronting parapsychologists is to decide whether MH is true or whether not-MH is true.

Exceptions to the general sensory-motor hypothesis SMH I shall call psi-phenomena, PSI. The hypothesis that such exceptions occur I shall call the PSI-hypothesis (PSIH). Again we have the logical equivalences:
1) SMH if, and only if, not-PSIH, and
2) PSIH if, and only if, not-SMH.
It follows that one of the most general tasks confronting parapsychology is to decide whether or not SMH is true (and, of course, formulate hypotheses about the relationship between exceptions to SMH and other phenomena, and the possible explanation of such exceptions).

ON THE PHILOSOPHICAL IMPORTANCE OF PARAPSYCHOLOGY

Granted that we understand parapsychology as the (validity-oriented) study of possible exceptions to the sensory hypothesis SH, the motor hypothesis MH, and the sensory-motor hypothesis SMH: What philosophical interest can such a study have? I shall not discuss all aspects of this question, but I shall attempt to say something about what general theoretical interest parapsychology can have even for people who don't think of themselves as especially interested in parapsychology. That is, I shall focus on some questions of general philosophical character that parapsychology exemplifies or illustrates in an interesting way.

First of all, it seems clear that parapsychology raises some conceptual questions which are of general interest. I'm here thinking of questions about how various parapsychological concepts are to be understood, such as: What is involved in the general concept of paranormal phenomena? The concept of extrasensory perception? Psychokinesis? It may look as if such concepts are so special that they hardly have any interest outside a parapsychological context. I take it that it is however of some general interest to understand the content of such general hypotheses as, say, the sensory hypothesis SH, and to understand the content of SH is the same as understanding the content of the negation of SH, which is the hypothesis ESPH. It follows that if it is of interest both to parapsychologists and non-parapsychologists to understand what is involved in a general hypothesis like SH, then it should also be of interest to both parapsychologists and non-parapsychologists to understand the content of the parapsychological hypothesis ESPH. In particular, the concepts which are involved in formulating a general hypothesis like SH are the very same concepts which are involved in formulating its negation, namely ESPH, so the concepts that we have to analyse in order to analyse the meaning of SH, are the very same concepts that we have to analyse in order to analyse the meaning of ESPH. Starting with the concepts of 'information' and 'sensory information', and from there on again to concepts involved in these concepts -e.g., the concepts of 'corresponds cq. does not correspond to actual states of affairs' or 'is due to chance cq. is not due to chance'. These are the very same concepts that we have to understand if we are to understand the content of the sensory hypothesis SH.

It also seems clear that parapsychology raises epistemological problems and problems in the philosophy of science that must be of general interest (note 3). Parapsychology will to some extent be distinct from other disciplines cq. scientific specialities through its choice of research objects (concentration on phenomena that seem to contain a psi-component, and so constitute an anomaly in relation to 'orthodox' theories of how human beings can interact with the environment), through its choice of perspective (it is a validity-oriented study of such seeming anomalies), and through its choice of research hypotheses (at least sometimes one accepts as a reasonable working hypothesis that the seeming anomalies are genuine anomalies). At the same time it should be clear from what has been said above that the concepts and methods that one finds within parapsychology must to a large extent overlap with the methods and concepts that one finds within other disciplines. Consider, e.g., the question of whether we have, in a certain case, an example of ESP. The question is whether or not we have the combination I & not-Se, i.e., information I without sensory information Se. That is, we must ask: Is there information I?

Is there a lack of (relevant) sensory information Se?
It is clear however that since the concept of information I is not peculiar to parapsychology, the methods and techniques that parapsychology has at its disposal for deciding the first question, are not peculiar to parapsychology either. On the contrary, in so far as parapsychology aspires to the status of a science, the methods at its disposal for deciding whether or not there is information, cannot be different from those we find in other sciences dealing with information. In the same way, just as the methods for deciding that there is sensory information are not peculiar to parapsychology, so the methods for deciding that there is no sensory information available are not peculiar to parapsychology. In this way the general methodological questions connected with the concepts involved in formulating the central parapsychological hypothesis ESPH, are of interest to any science making use of the concept of information and the concept of (direct or indirect) sensory information.

It is also clear that not only epistemological and methodological questions connected with the concepts used in formulating the central parapsychological hypotheses have general interest. Methodological and epistemological questions connected with the hypotheses themselves also have general interest. This follows from the fact that propositions about the epistemic status of the hypotheses ESPH, PK and PSIH (i.e., propositions about what counts in favour of, and what counts against, such hypotheses) are logically equivalent to propositions about the epistemic status of the hypotheses SH, MH and SMH. Consider, for instance, the following statements:
There are good reasons for believing that ESPH is true.
There are good reasons for believing that ESPH is false.
There aren't good reasons for believing that ESPH is true.
There aren't good reasons for believing that ESPH is false.
Such statements are logically equivalent to the following statements about the sensory hypothesis SH:
There are good reasons for believing that SH is false.
There are good reasons for believing that SH is true.
There aren't good reasons for believing that SH is false.
There aren't good reasons for believing that SH is true.

It follows that epistemic questions raised in connection with ESPH, are epistemic questions raised in connection with SH, just as epistemic questions raised in connection with SH are epistemic questions raised in connection with ESPH. Consider, for instance, the following questions about ESPH:

What kind of evidence would make it reasonable to accept ESPH? Can single, well-documented instances of ESP establish ESPH, or must we have reproducible instances of ESP before we can regard ESPH as established? Must we have a theory that can explain how ESP is possible before we can have good scientific reasons for accepting ESPH? What kind of evidence would make it reasonable to reject ESPH? How many apparent instances of ESP must we show to be only apparent instances of ESP, before we can conclude that ESPH is false?
Such questions can be reformulated to questions about SH:
What kind of evidence would make it reasonable to reject SH? Can single, well-documented exceptions to SH refute SH, or must we have reproducible exeptions to SH before we can regard SH as refuted? What kind of evidence makes it reasonable to accept SH? How many apparent exceptions to SH must we have shown are only apparent exceptions to SH before we can conclude that there are no such exceptions?

Assuming that epistemic and methodological questions connected with general assumptions like SH are of interest both to parapsychologists and non-parapsychologists, it follows that epistemic and methodological questions connected with ESPH are of interest both to parapsychologists and non-parapsychologists. Particularly significant, it seems to me, is the fact that in so far as paranormal phenomena are understood as anomalies or exeptions in relation to certain general hypotheses about how human beings are constrained in their interaction with the environment, the verification of such phenomena is the same as the falsification of those general hypotheses. To verify ESPH (showing that it is probably true, or reasonable to accept), is for instance the same thing as falsifying SH. So formulating questions about how ESP H, PKH and PSIH are to be verified is just another way of formulating questions about how to falsify general assumptions like SH, MH and SMH. That means that parapsychology in a striking way illustrates issues having to do with scientific rationality and falsification. An important problem within philosophy of science is that of how scientists behave, or ought to behave, when confronted with occurrences that seem to falsify their theories. E.g., how does one behave, or how ought one to behave, when confronted with seeming exceptions to SH? That is, how does one behave, or how ought one to behave, when confronted with seeming occurrences of ESP? Most scientists have remained sceptical about, or have adopted a wait-and-see attitude towards hypotheses like ESPH, PKH AND PSIH (note 4). That is another way of saying that they have remained sceptical about, or adopted a wait-and-see attitude towards claims to the effect that SH, MH or SMH have been falsified. The question then is what determines such attitudes, and how rational they are (note 5). Especially interesting in this connection is the now quite considerable body of literature critical of parapsychology, claiming that parapsychology has not succeeded in falsifying, or even cast serious doubt on, 'orthodox' principles such as the sensory-motor hypothesis (note 6). Such criticism may be looked on as expressing more or less legitimate (or illegitimate) strategies for avoiding the falsification of established theories like SMH.

PART II
THE PSI-HYPOTHESIS

TWO WAYS OF UNDERSTANDING THE SENSORY-MOTOR HYPOTHESIS

In this part I shall discuss in somewhat more detail what is involved in what I regard as the most general and important parapychological hypothesis, the psi-hypothesis PSIH. As explained earlier, such a discussion may, without loss of meaning, be reformulated as a discussion of the sensory-motor hypothesis SMH, and I shall make extensive use of this reformulation in the following remarks.

The sensory-motor hypothesis SMH says that non-accidental correspondences between (states and activities of) a subject and its environment come about through (direct or indirect) sensory interaction with the environment, and/or through (direct or indirect) motor interaction with the environment. I shall distinguish between two possible interpretations of this hypothesis: as a default principle, and as a universal principle.
As a default principle SMH says something about what is normally, or typically the case: normally, or typically, correspondences between subject and environment come about through the use of sensory-motor mechanisms.
As a universal principle SMH says that oe-correspondences always come about in this way: it excludes the idea that such correspondences, in so far as they are non-accidental correspondences, could have come about except through the use of sensory-motor mechanisms. In the same way SH and MH may either be understood as default principles or as universal principles.

Considered as default principles SMH, SH and MH articulate important aspects of our human experience. They articulate constraints on possible man-environment interactions that we take for granted in most contexts. We take it for granted that if we don't have guesswork, luck, coincidence, etc., then we have an adaptation to the environment that is either (directly or indirectly) the result of sensory information, or the result (directly or indirectly) of some motor influence on the environment. We can hinder people in getting information by hindering them in getting necessary sensory information (we hide something in a drawer, put it in an envelope, refuse to say something, etc.), and we can hinder people in influencing their environment by hindering their physical influence on the environment (we put something outside somebody's reach, we block somebody's movements, we lock somebody in, etc.). We have thus many, many times implicitly tested SMH, SH and MH and seen it confirmed that they express very real limitations on what non-accidental oe-correspondences are possible. Many institutions -lotteries, the secret police, jails, spies, etc.- could not have existed if hypotheses like SMH, SH and MH were not by and large valid. The fact that such institutions do exist may in itself be taken as evidence that such principles are at least valid as default principles.

On a deeper level the validity of SMH, SH and MH as default principles may be seen as reflecting our very existence as embodied beings. Part of what it means to say that a certain body B is the body of a person P, is that there are constraints on what objects in the environment P can have information about or influence, and that these constraints are centered around the body B. Things can for instance only be moved by me in so far as parts of my body can be moved by me, that is, in so far as there is this sort of limitation on what can be moved by me. In the same way, there are radical limitations on what I can be aware of (I cannot see the object that is hidden behind another object, and I'm not aware of it at all if I cannot infer its existence on the basis of something else that I perceive, etc.), and these limitations are centered around those parts of my body that we call sense organs. That it is parts of B that have this priviledged position is a presupposition for B being my body. In general, there cannot be that complex network of possibilities for action and awareness that characterise a living organism without a similarly complex network of impossibilities for action and awareness. SMH, SH, and MH considered as default principles may be taken as expressing such 'necessary impossibilities'.
Nothing of this excludes, however, that there are various forms of

exeptions to SMH. It only concerns SMH's validity as a default principle, i.e., as an expression of what is normally, or typically, the case.

It is only SMH as a universal principle that excludes the possibility that there are exceptions to it. One way of expressing this is to say that according to SMH, understood as a universal principle, all seeming exceptions to SMH can only be seeming exceptions, not real exceptions. This means that there are certain conjunctions of 'seemings' that must always include one or several 'seemings' that are only 'seemings'- i.e., some kind of illusions. Suppose, e.g., that there seems to be a nonaccidental correspondence between my description of a letter and the content of the letter. Then, according to SMH regarded as a universal principle, either it must a) only seems as if there is a non-accidental correspondence (i.e., it is really a question of guessing, being lucky, etc.), or b) it must only seem as if I have no (direct or indirect) sensory awareness of the content of the letter (I have read it earlier, or I have inferred what it is about, etc.), or c) it must only seem as if the description of the letter (or the thoughts behind the description of the letter) has no physically mediated influence on the content of the letter (I, or somebody else who has come to know what I believe about the letter, have changed the content of the letter, etc.).

Granted that it is reasonable to accept SMH (or SH, or MH) as some kind of default principle, is it also reasonable to accept SMH as a universal principle ?

Of course, if seeming exceptions to SMH never occur, this question is not of much interest. But it is an important aspect of human experience that it has a double structure. On the one hand, there are innumerable cases where things seem to happen in accordance with SMH, and where our actions and expectations are adapted to this fact. On the other hand, there is a comparatively small, but still significant number of cases where exceptions to SMH seem to occur. It is also a fact that the idea of such exceptions to SMH plays an important part in human life. The idea of such exceptions, conceptualised as e.g. prophesies, miracles, or possessions by the devil, plays an important part in human religious beliefs. It also plays an important part in beliefs in magic, witches, spells, etc. The seeming experience of, and belief in, exceptions to SMH also plays an important part in human life outside of such religious or superstitious contexts. Many people seem to experience, now and then, exceptions to SMH in the form of e.g. sudden premonitions, thought-readings, or dreams that come true. In fact, the experience of such seeming exceptions to SMH, and also the belief that some of them represent real exceptions to SMH, is fairly common, even in modern, industrialized cultures (note 7). It is therefore natural to ask, what reasons are there for accepting SMH, SH and MH not only as default principles, but also as universal principles, in spite of the many apparent exceptions to them?

SOME ASSUMPTIONS ABOUT EXCEPTIONS TO THE SENSORY-MOTOR HYPOTHESIS

In the following I shall only discuss SMH as a universal principle. The PSI-hypothesis PSIH is the negation of SMH understood in this way. I shall start by elaborating a point that is fairly trivial, but also fundamental. It is that SMH is really a fairly complex set of assumptions that it will always be possible to maintain in the face of any seeming exceptions to it. In other words, experience will never force us to abandon SMH, because by suitable modifications of one's assumptions one can make SMH fit any possible experience. This is not unique to SMH, of course, but it is no less true for SMH than for other theories. Some of the most important defensive strategies that SMH lends itself to, can be described in the following way:
We have an apparent PSI-event, i.e., an apparent exception to SMH, if we seem to have 1) an oe-correspondence C between subject and environment which 2) is not a chance coincidence A , and 3) is not the result (directly or indirectly) of sensory information Se, nor 4) a result , directly or indirectly, of motor influence Mo. So an apparent PSI-event is an event where we seem to have the following conjunction:

C & not-A & not-Se & not-Mo

If we believe in SMH and are confronted with an apparent PSI-event which tempts us to believe that PSI-events really occur, we have a dilemma. We cannot both accept SMH and accept that we are confronted with a genuine PSI-event. But of course, the conflict need not be resolved by discarding SMH. Instead one can maintain that what one is confronted with is only an apparent, and not a genuine PSI-event. That is, one may reject one or several of the conjuncts in the conjunction:

C & not-A & not-Se & not-Mo

The following table (where T stands for what one regards as true, and

F for what one regards as false) indicates the most important combinations that are possible:

	1	2	3	4	5
SMH	F	T	T	T	T
C	T	F	T	T	T
not-A	T	T	F	T	T
not-Se	T	T	T	F	T
not-Mo	T	T	T	T	F

1 stands for what we may call the paranormal option, namely that necessary and sufficient conditions for a PSI-event exist and that SMH is rejected. 2-5 indicate the most important alternatives to the paranormal option. Consider alternative 2. One can doubt that the supposed oe-correspondence really exists, and maintain that the person in question is wrongly described (perhaps because he himself remembers wrongly, or lies, or dramatizes, or exaggerates) as to the degree of correspondence between the states/activities of the person and events in the environment. Or consider alternative 3. Here one doubts that the oe-correspondence is a non-accidental correspondence. Since even very improbable coincidences now and then happen by chance alone, the claim that we really have an accidental coincidence A is of course impossible to refute definitely. Yet another alternative is 4. One may maintain that, in spite of appearances, we really have a case of direct or indirect sensory information. This is also an alternative which it is difficult to reject definitely, since there are so many ways in which one can have some form of indirect sensory information. In the same way it will be difficult to reject 5 definitely, i.e. that the oe-correspondence is the result, more or less directly or indirectly, of some physical influence on the environment. And of course, if it is difficult to reject definitely each of the alternatives 2-5 separately, it is even more difficult to refute definitely the indeterminate proposition that, somehow, one of the alternatives 2-5, rather than 1, must obtain.

In sum, there are so many loopholes when it comes to explaining away possible counterexamples to the sensory-motor hypothesis, that it is hard to imagine that anybody could be forced to give it up because of possible counterexamples (especially in a world where it is valid anyway as a default principle). This is perhaps one reason (among several) why it is accepted not only as a default principle, but also as a universal principle. In any case, if one is to seriously test SMH as a universal principle, one must not only (as pointed out above) accept what one may call the principle of normality:

PN: The sensory-motor hypothesis must in any case be valid as a default principle

but also what we may call the principle of fallibility:

PF: The sensory-motor hypothesis can under no circumstances be definitly falsified through counterexamples.

Reformulated as a proposition about PSI-events, PF can also be formulated thus:

PF: The PSI-hypothesis PSIH can under no circumstances be definitely verified by adducing counterexamples to SMH, i.e., through adducing examples of PSI-events.

But none of this excludes, of course, the possibility that there are cases of apparent PSI-events that in plausible and reasonable way can be explained within the framework of the sensory-motor hypothesis. Nor does it exclude the possibility that there are apparent counterexamples to the hypothesis that cannot be plausibly explained in this way. To a certain extent it is reasonable to expect that apparent exceptions to SMH can be explained away within the framework of SMH itself. Consider the matter from this point of view: To give a general account of how exceptions to SMH can be explained away within the framework of SMH itself (i.e., without making use of assumptions which are incompatible with SMH), involves showing how one would expect apparent exceptions to SMH to occur even within a universe where SMH was valid without exception. This one can do, along the following lines:

If there are apparent exceptions to SMH which are not real exceptions to SMH, this must be because, either
-there are apparent oe-correspondences C which are not real oe-correspondences C, or
-there are accidental correspondences A which only seem to be non-accidental correspondences, or
-there is an apparent lack of (direct or indirect) sensory information Se, which is not a real lack of sensory information, or
-there is an apparent absence of motor influence Mo which is not a real absence of motor influence.

This is a list of types of illusions that we know occur now and then. Hence we know that events must be expected that will give the false impression of being exceptions to SMH. We even know enough about such cases to give a fairly detailed account of how apparent exceptions to SMH are bound to occur. For instance, we know that human beings have problems judging the probability of certain coincidences happening by chance alone, and therefore may easily get the impression that something is a non-accidental coincidence when in fact it is an accidental one. We know that human beings may unconsciously remember things, or unconsciously infer things, or subliminally perceive things, so that one may get the impression that there is no relevant sensory information when there in fact is such information. We also know that human beings more or less unconsciously, and more or less indirectly, can physically influence their environment so that one gets the impression that there is no relevant motor influence when in fact there is motor influence (note 8). We know, then, that even if SMH is valid (as a universal principle), there must be events giving the impression of being exceptions to SMH. Such apparent exceptions to SMH are therefore of little importance as counterarguments to SMH. One can even maintain that SMH is supported not only by the events which obviously support it, but also by its apparent exceptions, because these seeming exceptions may be looked upon as necessary side-effects of just that cognitive and physiological apparatus that SMH presupposes. The following assumption therefore, which I shall call the principle of illusory exceptions, may also be regarded as a reasonable assumption to make when seriously testing the validity of SMH:

PIE: We must expect that many, probably most, apparent exceptions to the sensory-motor hypothesis SMH can be satisfactorily explained within the framework of that hypothesis itself.

If we call apparent PSI-events which are not real PSI-events for PSI-illusions, we can formulate this as a proposition about PSI, thus:

PIE: We must expect that many, probably most, apparent PSI- events are really PSI-illusions.

It is clear however, that even if many, perhaps most, apparent exceptions to SMH can be explained in this way, this is not necessarily true for all of them. If we are seriously to test SMH as a universal principle, we must actively seek for, or set up, situations where the sources of possible illusions mentioned above are eliminated, in order to see if exceptions to SMH still occur. If we are to seriously attempt to test the sensory-motor hypothesis therefore, the following assumption, which I shall call the test-principle, also seems to be a reasonable assumption:

TP: A serious attempt to falsify SMH presupposes that one actively seeks for, or sets up, situations where the possibilities for PSI-illusions are satisfactorily eliminated.

Formulated as a proposition about PSIH, we have:

TP: A serious attempt to verify the PSI-hypothesis PSIH presupposes that one actively seeks for, or sets up, situations where the possibilities for PSI-illusions are satisfactorily eliminated.

"Satisfactorily eliminated" does not mean "definitely eliminated" (cf. the principle of fallibility), but something like "what one has good reasons for believing eliminated". We are now back to parapsychology understood as the validity-oriented study of apparent PSI-events. If one is seriously interested in attempts to falsify the sensory-motor hypothesis, one must turn to parapsychology, since it is there that one finds most attempts to test SMH within the framework of what I call the test-principle TP, i.e., by actively seeking for, or setting up, situations where the possibilities for PSI-illusions are satisfactorily eliminated.

OUGHT ONE TO TAKE THE PSI-HYPOTHESIS SERIOUSLY?

The history of parapsychology could be regarded as repeated attempts to show that even when the possibility of PSI-illusions are satisfactorily eliminated, apparent exceptions to SMH will occur, and that consequently it is unreasonable to assume that all apparent exceptions to SMH can be explained within the framework of SMH itself, i.e., by only presupposing sensorimotor forms of interaction with the environment. No doubt many of these attempts have resulted in interesting negative findings, in the sense that they have documented that many apparent exceptions to SMH can be explained within the framework of SMH itself as being the result of misperception, methodological error, fraud, lapse of memory, etc. They have also often led to the result that when opportunities for PSI-illusions are satisfactorily eliminated, apparent exceptions to SMH have disappeared, something which might again be taken to indicate that apparent exceptions to SMH are just that-apparent exceptions (note 9). That one is able to explain away exceptions to a hypothesis like SMH within the framework of this hypothesis itself, is something that adds credibility to the hypothesis. The many negative research findings that one finds in the history of parapsychology are not just failures to find support for PSIH, they are findings that positively lend support to SMH. One can also maintain that these negative research findings indicate that the most fruitful research strategy for explaining exceptions to SMH, is to attempt to explain them within the framework of SMH itself.

The sensory-motor hypothesis is part of an extensive and elaborate system of theories that comprises not only SMH, SH, and MH, but also theories about the cognitive and physiological apparatus that makes possible our sensory-motor interactions with the environment. It is this whole corpus of knowledge, which also comprises theories about how different types of illusion are generated, that we can draw on when explaining apparent exceptions to SMH. The point isn't just that it is fruitful to attempt as far as possible to explain exceptions to SMH by use of known principles and mechanisms. One must of course attempt this, but the question is if it is always possible. One may maintain that the many negative results in parapsychology indicate that this in fact is possible, and that it will pay to direct one's efforts towards explaining exceptions to SMH in terms of those cognitive and physiological mechanisms that SMH itself is a reflection of.

It would however be seriously misleading to maintain that in all cases where the possibility of PSI-illusions has been satisfactorily eliminated, exceptions to SMH have failed to show. There are potentially three classes of parapsychological investigations: 1) Investigations where the possibility of PSI-illusions is satisfactorily eliminated, and where the result is negative, 2) investigations where the possibility of PSI-illusions is not satisfactorily eliminated, and where the result is either negative or positive, and 3) investigations where the possibility of PSI-illusions is satisfactorily eliminated, and where the result is positive in that exceptions to SMH still seem to occur. No doubt there are many parapsychological investigations that must be placed in either classes 1) or 2), but to maintain that all parapsychological investigations belong in either of these two classes, seems unreasonable. It seems to be the result of using one of the following two strategies: A) Using the definite elimination of possible PSI-illusions as a criterium for the satisfactory elimination of possible PSI-illusions, or B) using the fact that the investigation ends in a positive conclusion as a proof that possible PSI-illusions have not been satisfactorily eliminated. Since there is no reason to believe that SMH can be definitely falsified (cf. the principle of fallibility -the definite elimination of possible PSI-illusions in a parapsychological study with positive conclusion would be such a definite falsification), the use of strategy A) certainly means that all parapsychological studies must be placed in classes 1) or 2). The same goes for B), of course. It is also clear that if one operates with such criteria for the satisfactory elimination of possible PSI-illusions, SMH is made immune to possible falsification.

If such strategies that make SMH immune to falsification are abandoned, must we then conclude that SMH has in fact been refuted, that it has in fact been shown that it is more reasonable to reject then to accept SMH (as a universal principle)? Personally I'm not willing to go that far on the basis of the research material that I'm familiar with, since I also see (as many others, including parapsychologists, have) important reasons of a methodological and theoretical character which count against accepting the reasons in favour if PSIH as sufficient reasons for accepting PSIH (i.e., as sufficient reasons for rejecting SMH). I cannot go into the details of these arguments here. Let me however just briefly mention some of them.

1) To accept PSIH is nothing other than to reject the sensory-motor hypothesis SMH and all those theories about man-environment interaction which make use of SMH. It is not that one has any alternative to those theories.
That is, there is no alternative theory H which both implies that PSI-events will occur and overlap with SMH in the cases where it must still be considered valid. In that sense SMH and the physiological and cognitive theories connected with it, are the only usable theories we have at present. It is reasonable that one hesitates to give up the only usable theories one has.

2) In conformity with 1) one can also argue that we lack an important reason for accepting that exceptions to SMH really occur -what we can call theoretical reasons. That is, there is no theory H for which there exists independent empirical confirmation and which implies that exceptions to SMH occur. That one lacks such reasons may make one

doubt that the reasons we have for rejecting SMH (as a universal principle) are sufficient reasons for rejecting SMH.

3) One also finds wanting another type of reason for conceding that exceptions to SMH occur -what one may call experimentally repeatable reasons. That is, one has not succeeded in finding experimental conditions E such that exceptions to SMH reliably occur under those circumstances, so that also those sceptical about the occurrence of such phenomena may check for themselves that they really occur. This may again make one doubt that one has sufficient reasons for rejecting SMH.

4) The hypothesis SMH is not an isolated generalisation, but a generalisation connected with, and supported by other assumptions about the human cognitive apparatus. One knows, e.g., a great deal about how mental processes are dependent upon the brain which is again dependent upon nerve connections between brain and muscles and sense organs in order to influence the environment or receive information from it -something which in sum may be take as strongly supporting the view that a human being (as well as a human brain) is dependent upon sensorimotor mechanisms in order to obtain information from, and influence, the environment. It is this whole complex of assumptions surrounding SMH which is threatened when SMH is threatened. Let us call these assumptions H. We then have, schematically: If PSIH, then not-H, i.e., we have: If H, then not-PSIH. Consequently, in so far as there is empirical support for H, there is also empirical support for not-PSIH, i.e., for the assumption that exceptions to SMH do not occur.

5) Various seemingly well-documented exceptions to SMH which have been reported in parapsychology have apparently not only been exceptions to SMH and physiological and psychological principles connected with it, but also exceptions to certain general principles which concern not only psychology and physiology, but also physics and chemistry, and in some cases (one could argue) all science. The most striking example is that some exceptions to the sensory hypothesis SH seem to involve not only present events, but also future events. If we assume that the information in question presupposes a causal connection (in the widest sense) going from what one has information about to the states and activities of the subject where this information find its expression (dreams, guesses, premonitions, etc.), then we don't only have an exception to SMH, but also an exception to the general principle that future event cannot influence past events. Let us call this principle

the causal-temporal principle CT, and the hypothesis that exceptions to SMH which are also exceptions to CT really do occur, the precognitive hypothesis PCH. Then we have: If PCH, then not-CT, i.e., if CT then not-PCH. So, if there are good reasons for assuming CT (empirical reasons, and perhaps also conceptual or metaphysical reasons), then there are also good reasons for assuming that at least the type of exceptions to SMH which are precognitive events, do not really occur.

1)-5) indicate that we do not yet have sufficient reason for maintaining PSIH, i.e., for rejecting the sensory-motor hypothesis. One thing is however to lack sufficient reason for preferring not-SMH (i.e., PSIH) to SMH, another is to maintain that one has no reasons -significant and important reasons- for doubting SMH. It seems to me that the apparent exceptions to SMH that we know from parapsychology are sufficient to cast serious doubt on the proposition that exceptions to the sensory-motor hypothesis never occur. In so far as this is the case there are good reasons for taking the PSI-hypothesis PSIH seriously. There are also good reasons for devoting time and resources to examining it. Our reasons for continuing the investigation of a hypothesis is not that we know it to be true, but that we have interesting reasons for believing that it may be true.

NOTES

1) Partly as a result of what is sometimes referred to as post-empiricist philosophy of science (i.e., the criticism, found in the works of Feyerabend, Hanson, Kuhn, Laudan, Polanyi, Toulmin, of logical empiricism and its ahistorical approach to the philosophy of science), the demarcation between science and pseudo-science has become increasingly problematic. This may be one reason why parapsychology has increasingly attracted the attention of philosophers, historians and sociologists of science. Anthologies and monographs illustrating this trend are Cerullo 1982, Collins & Pinch 1982, Grim 1982, Hanen, Osler & Weyant 1980, Laudan 1983, Mauskopf 1979, Mauskopf & Vaugh 1980, McClenon 1984, Wallis 1979.

2) Summing up his review of philosophers' contributions to parapsychological research, J.F.Nicol writes: "If we consider the mere quantity of work published by philosophers, we need to realize that philosophy is a very small profession ... it is safe to say that

philosophers have produced more research than other professions" (Nicol 1976, p.168)
The list of well known philosophers who have functioned as presidents of The Society for Psychical Research is impressive: H.Sidgwick; W.James; H.Bergson; F.C.S.Schiller; H.Driesch; C.D.Broad; H.H.Price; C.W.K.Mundle.
Philosophers' interest in philosophical questions connected with parapsychology are exemplified by several anthologies, e.g., Ludwig 1978, Shapin & Coly 1976, Thakur 1976, Wheatley & Edge 1976. Among well known philosophers who have made a quite considerable contribution to the literature connected with parapsychology are C.D.Broad, C.J.Ducasse and H.H.Price. They are all represented in Wheatly & Edge 1976.

3) To philosophers the sensory hypothesis SH will be of particular interest in that it can be interpreted as one way in which to formulate the empiricist principle that all knowledge (of a non-analytical character) must build on sense experience. What is interesting with SH in a parapsychological context is that it is treated as an empirical thesis which can be either undermined or supported by empirical evidence. A principle like SH can of course be interpreted in many different ways, and not all interpretations will make it a hypothesis that can be tested empirically. If one is of the opinion (see e.g. Quine 'Epistemology naturalized') that an empiricist principle ought to be interpreted as some kind of empirical thesis, then this way of interpreting SH should be of particular interest. But even if one doesn't agree with this position, it should still be of interest to see elucidated, conceptually and epistemologically, what is involved in a sensory hypothesis SH when it is interpreted as some kind of empirical proposition. This discussion will then coincide with the parapsychological-philosophical discussion of what is involved in the concept of ESP, i.e., what is involved in the concept of exceptions to SH.

4) In the literature mentioned in note 1, one will find many examples of seemingly positive findings in parapsychology being fiercely challenged. We have to remind ourselves, however, that is may be difficult to judge how intense the resistance, or scepticism, towards parapsychology is among scientists in general. One thing is the attitudes of the most engaged combatants, another is the attitudes of scientists in general. Doubt and uncertainty, rather than dogmatism and aggressive certainty, seem to have been typical of the famous controversies in the history of parapsychology (cf, Mauskopf & Vaugh

1976:'Parapsychology and the American Psychologists: A study of scientific ambivalence';and Palfreman 1979:'Between scepticism and credulity: A study of victorian scientific attitudes to modern spiritualism').
Scepticism about the truth of a proposition is of course compatible with a positive interest in seeing the issue examined further. In surveying members of the American Psychological Association, Warner and Clark found that only 9% had a positive attitude as to the probability of ESP really existing, but 89% were of the opinion that ESP was a legitimate field of research (Warner & Clark 1938). In another survey of APA-members (Warner 1952) 17% had a positive attitude towards the probability of ESP existing, but again 89% had a positive attitude towards ESP as a legitimate field of research.

5) To what extent it will appear rational to uphold to the sensory-motor hypothesis SMH -that is, reject the PSI-hypothesis PSIH- in spite of seeming counterexamples (i.e., seeming examples of PSI), may partly depend upon fundamental methodological assumptions. There are several possibilities. E.g., if one sides with Popper in emphasizing the importance of repeatable counterexamples to a general hypothesis as essential to its falsification (c.f. Popper 1974, p.86), then the lack of repeatable experimental results that one typically finds in parapsychological research, may be a decisive reason for upholding SMH in spite of seeming counter-examples. If one sides with Kuhn (c.f. Kuhn 1970, p.24) and stresses the functional importance for the development of a normal science (i.e., the elaboration of a paradigm) that one not give up a paradigm too quickly in the face of seeming anomalies, then it may also appear rational to uphold SMH (considered as part of a paradigm) in spite of apparent exeptions to it. Inspired by Lakatos (c.f. Lakatos 1980, p.118), one might also regard SMH as part of the hard core of a research program for exploring the interaction between man and environment, and regard the PSI-hypothesis PSIH as part of (or the germ of) another such research program. If one then looks at SMH as part of a still progressive research program which (contrary to research programs incorporating the PSI-hypothesis) are producing interesting results (e.g. when it comes to understanding various forms of apparently paranormal phenomena), it may also appear rational to uphold SMH in spite of exceptions to it which resist a 'normal' interpretation. For my own part I have no problems accepting that there are reasons why one should not reject SMH (cf. 6). I have, however, some problems seeing why it may not also be worth while, in the light of parapsychological research findings, to explore another possiblity at the same time,

namely that the PSI-hypothesis is in fact correct, and that we may in the end have to find some alternative to SMH (considered as a universal principle) and theories incorporating SMH.

6) Criticism of parapsychology can be found in e.g., Alcock 1981, Marks & Kammann 1980, Hansel 1980, Kurtz 1985, and Taylor 1980. In these books one find exemplified all the strategies for defending the sensory-motor hypothesis that I sketch in 5). Different authors put different weight on different strategies. Marks & Kammann 1980, emphasizes the problems human beings have in judging the probability of chance coincidences (so that one sees unexplainable connections where there really are no connections). In Hansel 1980 the possibility of direct or indirect sensory information/motor manipulation even in seemingly strictly controlled experimental conditions, plays a more important part. The weight that Hansel and several other critics of parapsychology (particularly G.Price 1955) have put on possible frauds when explaining away possible exceptions to SMH is remarkable. One can of course always save the sensory-motor hypothesis from falsification by assuming sufficiently complicated cases of fraud. In some (but by no means all) their attacks on parapsychology Hansel, and several other critics of parapsychology, have come dangerously close to making the sensory-motor hypothesis immune to any possible falsification.

7) Systematic surveys seem to confirm this. E.g., in Danish survey from 1957 about one out of ten claims to have had some psychic experience. In a German survey about one out of five makes a similar claim. In an Icelandic survey from 1974 (a country which seem to have strong 'psychic' traditions) the number is as high as two in three (for comparisons, and details of the Icelandic survey, see Haraldsson et al. 1976). Such surveys are of course to be treated with care, but at least they seem to indicate that experiences that give the subjective impression of involving a psi-component, are quite frequent. Not only the experience of apparently paranormal events seems to be fairly frequent, the belief that such events really occur also seems to be frequent. After a review of some of the more important surveys in this field, Alcock (Alcock 1981) concludes: "... belief in the paranormal, psychologists aside, is currently very common. Various studies of university undergraduate and graduate students . . have indicated that amount of university education has virtually no effect on this" (p.27).

It is of some interest that persons with an academic background often seem to believe that e.g. ESP really occurs. One of the papers that Alcock refers to is "Attitudes of college professors toward extrasensory perception" from 1970 (U.S.A.). Here Wagner and Monet report that 66% of those responding (1184 out of 2400) were disposed to believe that ESP could occur, while 23% were sceptical. Wagner and Monet compared this with the attitudes toward paranormal phenomena expressed in a Gallup poll, where only one-half of the sample expressed belief in ESP, and concluded that their group had attitudes towards ESP that were more positive that those of the public in general. Again we have an indication that scepticism towards the sensory-motor hypothesis as an exceptionless principle is quite frequent, even among people trained in 'established' or 'orthodox' academic disciplines.

8. Both Marks and Kammann 1980 (see especially chap. 11 and 12) and Alcock 1981 (see especially chap. 4 and 5) have interesting discussions of various 'normal' cognitive mechanisms that may result in (subjective) paranormal experiences. See also Morris 1986.

9. Some of the institutions of parapsychology (e.g. the Society for Psychical Research) have done much useful work in this respect. Getting a correct picture of how often parapsychological experiments give a negative (nonsignificant) result, is notoriously difficult, because many negative results are never published. Judging from published material found in parapsychological research journals, the yearly 'Research in Parapsychology', and reviews like e.g. Palmer 1980, negative results must be quite common.

ABSTRACT

This paper is in two parts. In part one I discuss the concept of paranormal phenomena and what general theoretical and philosophical interest the study of such phenomena may have. Parapsychology is the validity-oriented study of possible exceptions to certain general and rather complex principles of how human beings (or other organisms) can obtain information about, or influence, their environment. Parapsychology is of general theoretical and philosophical interest in so far as the study of these principles is of general theoretical and philosophical interest.

In part two I discuss in somewhat more detail what is involved in the study of possible psi-phenomena. Psi-phenomena are understood to be exceptions to the general principle that we can only obtain information about, or influence, the environment through sensory or motor interaction (the sensory-motor hypothesis). It is also discussed to what extent one ought to take seriously the hypothesis that psi-events really occur.

REFERENCES

Alcock, J.E. 'Parapsychology: Science or magic'. Toronto, 1981.

Allison, P.D. 'Experimental parapsychology as a rejected science', in Wallis, 1979.

Cerullo, J.J. 'The secularization of the soul. Psychical research in modern Britain', Phil., 1982.

Collins, H.M. and Pinch, T.J. 'Frames of meaning. The social construction of extraordinary science', London, 1982.

Collins, H.M. and Pinch, T.J. 'The construction of the paranormal: Nothing unscientific is happening', in Wallis, 1979.

Edge, H.L., Morris, R.L., Palmer, J. and Rush, J.H. 'Foundations of parapsychology', London, 1986.

Grim, P. (ed.) 'Philosophy of science and the occult', New York, 1982.

Hanen, M.P., Osler, M.J., Weyant, R.G. (eds.) 'Science, pseudo-science and society', Wilfrid Laurier University Press, 1980.

Hansel, C.E.M. 'ESP and parapsychology: A critical re-evaluation', New York, 1980.

Haraldsson, E., Gudmundsdottir, A., Ragnarsson, A., Loftsson, J., Jonsson, S, 'National survey of psychical experiences and attitudes towards the paranormal in Iceland', in Research in Parapsychology 1976, London, 1977.

Johnson, M. 'Parapsykologi', Uddevalla 1980.

Krippner, S (ed.) 'Advances in parapsychological research. Part II. Extrasensory perception', New York, 1978.

Kuhn, T.S. 'The structure of scientific revolutions', Chicago, 1970.

Kurtz, P. (ed.) 'A sceptic's handbook of parapsychology', Buffalo, 1984.

Lakatos, I.M. and Musgrave, A (eds.) 'Criticism and the growth of knowledge', London, 1980.

Lakatos, I.M. 'Falsification and the methodology of scientific research programs', in Lakatos and Musgrave, 1980.

Laudan, R. (ed.) 'The demarcation between science and pseudoscience', Working papers in science and technology, Volume 2, Blacksburg, Virginia, 1983.

Ludwig, J. (ed.) 'Philosophy and parapsychology', New York, 1978.

Mauskopf, S.H. (ed.) 'The reception of unconventional science', A.A.A.S. Selected Symposium 25, Boulder 1979.

Mauskopf, S.H. and Waugh, M.R. 'Parapsychology and the American psychologists: A study of scientific ambivalence', in Shapin and Coly, 1976.

Mauskopf, S.H. and Waugh, M.R. 'The elusive science', J. Hopkins University Press, 1980.

Marks, D. and Kammann, R. 'The psychology of the psychic', New York, 1980.

McClenon, J. 'Deviant science: The case of parapsychology', Philadelphia, 1984.

Morris, R.L. 'What psi is not: The necessity for experiments', in Edge, Morris, Palmer and Rush, 1986.

Nicol, F.J. 'Philosophers as psychic investigators', in Shapin and Coly, 1976.

Palmer, J. 'Extrasensory perception: Research findings', in Krippner, 1978.

Palfreman, J. 'Between scepticism and credulity: A study of Victorian scientific attitudes to modern spiritualism', in Wallis, 1979.

Popper, K.R. 'The logic of scientific discovery', London, 1974.

Price, G.R. 'Science and the supernatural', Science, 122, 1955, 359-367.

Quine, W.V.O. 'Epistemology naturalized', in Quine, 1969.

Quine, W.V.O. 'Ontological relativity and other essays', London, 1969.

Shapin, B. and Coly, L. (eds.) 'The philosophy of parapsychology', New York, 1976.

Thakur, S.C. (ed.) 'Philosophy and psychical research', London, 1976.

Wagner, H.H. and Monnet, M. 'Attitudes of college professors toward extrasensory perception', Zetetic Scholar, 5, 1979, 7-16.

Wallis, R. (ed.) 'On the margins of science', Sociological Review Monograph 27, University of Keel, 1979.

Warner, L. 'A second survey of psychological opinion on ESP', J.o.P., 16, 1952, 284-295.

Warner, L. and Clark, C. 'A survey of psychological opinion on ESP', J.o.P, 2, 1938, 296-301

Wheatly, J.M.O. and Edge, H.L. (eds.) 'Philosophical dimensions of parapsychology', Illinois 1976.

Magne Dybvig,
The University of Trondheim,
Department of Philosophy,
N-7055 Dragvoll,
Norway

EUR.J.PARAPSYCHOL., V.6, 1986

LETTERS TO THE EDITOR

From Werner Eeman:

With regard to Haraldsson's and Gissurarsson's article 'Perceptual defensiveness, Ganzfeld and the percipient-order effect' (November 1985 issue) I would like to make the following comment: In the section "Results' of their pilot experiment the authors state that their 'percipient-order effect' seemed to be confirmed as "in 16 sessions the first percipients obtained 2 hits, whereas the second percipients obtained 7 hits in an equal number of sessions'.

Is this the only interpretation for the results obtained in the pilot experiment? Consider for example the possibility of a (conscious or subconscious) strategy used by the percipients when "thinking out loud" their impressions during the Ganzfeld session: When they were suppossed to score 'negatively', they could have concentrated on just one item and talked exclusively about it. In this case the chance of obtaining direct hits was very small. If, on the other hand, they had to concentrate on obtaining as much hits as possible, they could have mentioned anything they had in mind, hoping that one of their 'guesses' corresponded with some aspect of the target picture. Of course I do not pretend that this is what happend in Haraldsson's and Gissurarsson's pilot experiment, but it seems to be a possibility that one has to be aware of when evaluating a free-response GESP-experiment.

Werner Eeman
Meirveld 28
9310 Lede
Belgium

CALL FOR PAPERS FOR THE 30th ANNUAL CONVENTION OF THE PARAPSYCHOLOGICAL ASSOCIATION

The 30th annual convention of the Parapsychological Association will be held Tuesday, August 4 to Saturday, August 8, 1987 at the University of Edinburgh. Persons interested in attending the convention may write to the Arrangements Chairperson, Dr. John Beloff, Psychology Department, 7 George Square, Edinburgh EH8 9JZ, Scotland.

Anyone may submit a paper or a poster for consideration by the Program Committee. Papers may be empirical, theoretical, or methodological in nature, but the Program Committee will not consider papers published elsewhere prior to the convention.

Papers should adhere to the style of the 'Publication Manual of the American Psychological Association' (3rd edition), or the format used by the British Journal of Psychology. They should be typed on one side of 8.5 x 11 or A4 paper, and each page should have minimum margins of one (1) inch on all sides. The first sheet should have a centered title, author(s), and affiliation, followed by an abstract of no more than 300 words.

Text should not exceed 12 single-spaced pages (6000 words), with no more than five additional pages for essential figures, tables, and references. If possible (e.g., using a word-processor) prepare and submit one single-spaced copy and three double-spaced copies; otherwise send four single-spaced copies. All submissions must be carefully edited and finished in camera ready form for inclusion in the Convention Proceedings. All your copies of proposed papers must be received by the Program Chairman by the deadline of April 30, 1987.

Presentation time will be 20 to 30 minutes, including a question period. Indicate exactly what audio-visual aid you will need, and indicate which of multiple authors will make the presentation. In absentia presentations will be allowed only in exceptional circumstances.

Posters are brief papers or other materials presented on poster board in an installation separate from the convention floor. Proposals for posters must include four copies of all material to be presented in the poster and an estimate of the size of the required posters. Photocopies of photographs are acceptable. This material must reach the Program Chairperson by May 31, 1987.

Members and associates only of the Parapsychological Association may propose symposia, panel discussions and workshops. Symposia are formal presentations by participants on related topics. Proposals for symposia must include four copies of a summary sheet indicating title, chairperson, participants, order of presentation, and proposed time allotments, up to a total of 90 minutes, including discussion periods. Proposals must also include a full paper, prepared as detailed above, from each participant. This complete package must reach the Program Chairperson by April 30, 1987.

Panel Discussions are informal round-table discussions intended to maximize spontaneous interactions of participants and the audience. Formal presentations should not exceed five minutes. Proposals for panel discussions must include four copies of a summary sheet including a title, chairperson, participants, order of presentation, and time allotments up to a total of 90 minutes, and an abstract of less than 500 words from each panelist. The complete package must reach the Program Chairperson by April 30, 1987.

Workshops are informal discussions of specific topics. Proposals for workshops should provide a title, chairperson, participants, and description of workshop activity. Workshops will not be listed as part of the formal convention program, but will be announced during the convention. Workshop proposals will be accepted until June 15, 1987.

Address all correspondence regarding the program to:

Professor Robert Morris, Chairperson
Parapsychological Association Program Committee
Psychology Department
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ, Scotland

EUR.J.PARAPSYCHOL., V.6, 1986

RNG-PK MICROCOMPUTER 'GAMES' OVERVIEWED: AN EXPERIMENT WITH THE VIDEOGAME 'PSI INVADERS'

Loftur Reimar Gissurarson

Most of the early parapsychologists used specially gifted subjects in their experiments, but since then, the emphasis has been on testing unselected subjects. This change in emphasis has given rise to the 'experimenter effect' amongst other effects and controversies.Psychokinesis (PK), which seems to follow the same tracks as other PSI abilities regarding various effects, has been defined as the extramotor aspect of psi, being direct influence (meaning mental but nonmuscular) exerted by the participant on an external physical process, condition, or object. Today parapsychologists tend to make a distinction between macro and micro PK, macro being phenomena like metal-bending or gross movements of objects e.g. poltergeists and phenomena of physical mediumship, while micro being small scale phenomena, detectable by statistically significant deviation from mean chance expectation (MCE) at the 5% criterion level ($p<.05$).

Computer games are considered to have advantages over old fashioned methods in measuring PSI (see for an overview Brougthon 1981; 1982). Firstly, computerizing a game automates some parts of the experiment, e.g. freeing the experimenter from data recording duties -the data collection being automatically stored and obviating the need for a double blind setup, since the computer design serves as a substitute for the other experimenter. A psi computer game therefore makes the whole experiment easier and the experimenter can be freed while it is running. Secondly, computers have the advantage of not making mistakes and are less sensitive to biases, recording errors or fraud on the participant's behalf. Thirdly, if one is already using a computer to run an experiment, additional programs can be incorporated to do the statistical work. Having the random number generator (RNG), as the element of randomizing the target, rules out nonrandomness in the target sequence with reasonable confidence. To exclude uncertainty, simultaneous control condition might be used to accomplish checks on the computer and the randomness. Finally, computer games might elicit fun or excitement and in that way perhaps promote an environment which may produce greater evidence of psi. This has been the main argument for their usage.

A RNG-PK microcomputer controlled video game, called PSI INVADERS, was used to evoke PK performance from 15 subjects tested in the U.S. and 15 subjects tested in Iceland. Each participant played individually one warming-up and two 'real' games. The purpose was to observe how effective this allegedly motivating game, PSI INVADERS, is in producing evidence for PK in two culturally different groups. This is the first published article on results from PSI INVADERS, although the game has been in use for some time.

The authors of the program package PsiLab, which contains PSI INVADERS (Berger and Honorton 1985; for details see the Manual for PsiLab 1985), hoped that it would facilitate increased inter-experimenter replication rates by providing a uniform environment for systematic replication through standardized data collection and analysis procedures, and the use of similar hardware with known operating characteristics (Berger and Honorton 1985, p.71). The experiment which is to be reported is thought of being an answer to that plea.

Overview on published RNG-PK computer 'games'

The overview is restricted to purportedly motivating game-like RNG-PK computer setups to be used with unselected subjects, games that make extensive use of high-resolution graphics and sound effects. There are many computer controlled RNG-PK setups in use, but they are not of concern here since they cannot be considered 'games' that per se are supposed to enhance interest and motivation, e.g. the BASIC computer program, ALGERNON, which provides meaningful information and answers to important personal queries if psi is manifested (Braud and Schroeter, 1983). Based on the same principle is the PK version of the computer game, ALICE, its program being written in BASIC (Varvoglis and McCarthy 1986). In above two games the subjects are expected to psychokinetically influence the answer-selection process, which rather surprises and entertains the subjects. Furthermore, one can mention the software program HORIZON (Debes and Morris 1982; Talbert and Debes 1982) ,which is a computer screen version of the Rhinean PK technique of dice throwing for positions referred to as a placement method/test of PK (Rhine and Pratt 1957, p. 153-155).

Brougton (1979) has traced predecessors of the high-technology computer games to experiments done by Steen (1957), Ratte (1960) and Ratte and Greene (1960). Steen for instance, incorporated a PK test into a simulated baseball game played with a dice. Ratte and Greene used a basket ball style game with a dice.

Ratte (1960) used a dice game called PK-Basketball, a standard dice testing technique in which the participant in competition with another player, threw two dice for a total of seven. Ratte´s general results favored the gaming technique over a noncompetitive situation but the differences were insignificant.

Some reports about RNG computer games measuring PK have been published in Research in Parapsychology (RIP) and the Journal of Parapsychology, but very few as articles. First official reports of experiments using ´motivating´ PK computer games appeared around 1978, from exploratory experiments conducted by Weiner (1978) and Beloff, Broughton and Wilson (1978). Beloff et al. (1978), did an informal experiment, using a computer game, in which the PK element was embedded. It is unfortunate that only a few games were conducted, and that no results are to be found.

Weiner (1978) used a computer game simulating a horse race, where the computer assigned ´bets´ of low risk, $5, and of high risk, $25. The game was based on 4 columns of numbers, ´horses´, that counted upward from zero and were displayed on the computer terminal. A random number generator (p=.5) determined whether or not the numbers in the various columns proceeded to the next higher numbers so that at the end of 50 trials ´race´ the four columns showed different counts. Subjects choose one of the ´horses´ and tried to cause it to finish the race with the highest count. Weiner hypothesized that the 19 participants tested, would score better on the high-risk than on the low-risk race. None of the three major analyses (comparing the final counts of the high- and low-risk races and of each condition with chance) were significant. Secondary analysis showed that the 11 participants tested individually scored much higher than the 8 tested in groups.

Weiner (1979) used a computer game where participants used their motor skills to manipulate a dial in order to keep a bar centered on the screen. The difficulty of the task increased in steps until the bar became so unstable that it moved off the screen. A PK test was incorporated into this test by an RNG interface, such that PK hits would help the participant to keep the bar under control by preventing the task from incrementing to greater difficulties. The results of the relationship between PK and motor skill of 26 subjects were insignificant.

Broughton published in 1979 his results of an experiment, in which he utilised a PK computer game, called THE HEAD OF JUT (Broughton, 1979). Broughton used a device, on which was mounted a column of 32 small red lamps. Alongside the column of ligths was a slot for a strip of paper which bore labels for different points on the scale. Lamp number 25 was labeled ´average´ and this represented the MCE for a run. Other labels ranged from 17 ´terrible´ and above the 32nd light ´outstanding´. On top of the device, over the column, was a bell which rang with a single ´ding´ whenever the score exceeded 32. The participant initiated a run by pressing a button mounted in the base. The lights would begin lighting from the bottom upwards, very rapidly at first but with decreasing speed, as it came to a stop, and then fell back. The device was controlled by a computer. Broughton used two groups of participants, those who played individually and those who played among friends. The latter condition was expected to foster playful competition. Split analysis technique shifted the data in pilot and confirmatory design. Pilot PK data showed no significance departures from chance in general. The confirmatory PK data failed also to demonstrate PK effects.

Honorton, Barker and Sondow (1983) conducted a microcomputer-based RNG study comparing immediate versus delayed feedback on RNG hit rates. Three experimental series were done. Statistically significant overall RNG effect combining feedback and silent binary hits for only series 1, were predicted. Event-by-event feedback was available to subjects in the feedback condition. Feedback to the silent source, was limited to an end-of-trial statistical summary. The feedback source was displayed to participants via a thermometer-style computer graphics display showing a bar rising and falling in relation to the current feedback source byte value. Centered horizontal lines on either side of the bar demarcated target/nontarget areas of the display. Arrows on either side of the bar displayed the vertical target location. Bar colour provided feedback on cumulative performance within the trial, white bar was associated with scoring above chance, red with scoring below chance, and yellow at chance. A special 'Jackpot' display was activated at the end of the trial if a preset scoring threshold was met. A total of 310 control trials were taken from each RNG source at various points during the experimental series. No overall significant scoring was found in any of the three series.

Schechter, Barker and Varvoglis (1983) combined a computer controlled video game, PSI BALL, with a RNG-PK task. The player moved a lever to keep a 'ball' on the TV screen away from the screen's 'walls' for as long as possible. About five times each second, a ten-event RNG trial was taken. If there were fewer than five hits in the trial, the game's difficulty was increased by making the ball slightly more sensitive to small lever movements. The difficulty did not change if five or more of the RNG events were hits. Ten players participated. The game's difficulty level was affected with RNG hits in half of 10 sessions, but hits did not affect the game in the other half (referred to as 'contingent' RNG runs since scoring on them has effect on the game, opposed to 'noncontingent' where RNG runs do not affect the game). There were no significant differences among conditions or deviations from MCE.

Schechter, Barker and Varvoglis (1984) reported results of another study with PSI BALL. This second pilot study involved four slight changes from the pilot 1 procedures, mainly on the statistics. As in pilot 1, the four sets of ten million RNG events each that were gathered, as a check on the RNG's general performance, showed no significant departures from chance performance. The only aspect of the data suggesting psi, was that the RNG z-scores were positively and significantly correlated with the mean number of RNG trials per session in the contingent games. RNG performance was not significantly related to any of the other measures. Schechter et al. concluded that "the results of the two pilot studies suggest that if the current version of PSI BALL is eliciting any PSI at all, it is too weak an effect to be useful for process-oriented research with small samples".

Schechter, Honorton, Barker and Varvoglis (1984) reported an experiment they conducted on relationship between RNG scores on two computer controlled RNG-PK games, and some of the participants' psychological characteristics. In the game, VOLITION, the participants received immediate trial-by-trial feedback which included both auditory as well as visual components for directional performance, but in PSI BALL there was no such feedback. In VOLITION the RNG feedback display was presented as an on-line graph of the cumulative deviation from chance, providing the player with a clear picture of cumulative performance relative to goal. The computer-graphics display showed zones of significance, the developing cumulative deviation line, and a variety of audio/visual rewards for individually strong trial scores analogous to the 'difficulty level' options on commercial video games. The player's task was to produce high or low RNG values (above or below chance) according to High or Low aim, displayed and selected at the start of the game. Each game consisted of 100 trials and 100 RNG events each. Participants' psychological characteristics were measured with two questionnaires. Relationship between RNG and psychological characteristic scores of the 10 participants, and between VOLITION scores and PSI BALL scores were compared but no significant relationships were found.

Broughton and Perlstrom (1985a) modified an existing commercially available APPLE II microcomputer game called OINK (Beagle Bros., Inc.). The game consisted of a number of 'turns' of five rolls of a pair of dice displayed on the computer screen. The player's goal was to obtain as high a number as possible. If a double was obtained, that score was not counted and all points accumulated up to that point in the turn were erased. Major modifications were made to the original program such that it became the PK test, P-OINK, with embedded hardware RNG. Participants were able to play with the 'computer' and the game terminated when player or computer exceeded 200 points at the end of his fixed turn of five rolls of the pair of dice. In examining the performance of single participants, in contrast to 'joint' participation in the P-OINK game (subjects led into believing they were competing with an unseen player linked by a telephone hookup), the result of the PK component of the game did not differ significantly from chance for the 50 participants. In a competitive condition the participants' scores were suggestively high but near chance in a noncompetitive condition. The difference between the two conditions was insignificant.

Broughton and Perlstrom (1985b) again reported an experiment using a micro-PK test in the form of the competitive game, P-OINK. None of the results of the game scores differed significantly from MCE. They, however, found a significant negative correlation between participants scores and state anxiety, $r=-.49$ (26 df, $p<.01$, one-tailed),which confirmed their post-hoc findings in the earlier experiment. Broughton and Perlstrom concluded that their two experiments demonstrated a negative relationship between PK performance and levels of anxiety at the time of performance, i.e., participants who reported themselves to be more anxious obtained lower scores in the PK test.

The Computer Game: PSI INVADERS

The RNG based APPLE computer game, PSI INVADERS, is a software package included in 'PsiLab', which is a computer hardware/software system for psi researchers with APPLE series computers. It was produced and developed by the Psychophysical Research Laboratories (PRL) in Princeton N.J. (for details see the Manual for PsiLab 1985; Berger and Honorton 1985). PSI INVADERS is an adaptation of the popular arcade game SPACE INVADERS and can be presented or conceived as either PK or ESP task. In present experiment, the participants were instructed to think of the game as a PK task, i.e. to try with each press of a button to 'affect' the RNG.

The purpose of the game for the participant is to shoot down invaders from space with a laser gun, while trying to avoid being hit by them (all information and instructions appear on the screen before the game begins). Players are to press a button on a game paddle to fire their laser. The cover story is, that the laser gun is old and frequently misfires, so the player is asked to use 'The Force' to make it fire.

Laser firing is contingent upon the output of the RNG. With each press of the game paddle, the RNG is sampled one 'run'. Each run consists of 100 binary trials (where $p=.5$). For each trial a bit from the RNG is compared to a target bit which alternates between 0 and 1, thus avoiding bias in the RNG. If the RNG sample bit and the target bit are the same, the trial is counted as a hit. Run scores of 51 or greater are required for the laser to fire. Run scores of 50 or less result in a 'misfire'.

In all the participant has 100 button presses but it is, of course highly unlikely that the gun ´fires´ on each button press (the run score MCE=50; sd=5). Whether the gun fires depends, viz., upon the participants PK. When the 100 button presses are finished, the game is over. Actually, the only component measured is how well the participant manages to make the randomly shooting gun - fire. The invaders can not destroy the gun, and all game points that are given (displayed at the bottom center of the game screen; for firing the laser, subtraction from points for getting hits and so forth) are supposed to motivate the participant to fight. These game scores do not correlate directly with the z-scores. Hence, it is possible for a player to receive large game points and have a chance, or below chance, z-score.

Furthermore, trial-by-trial feedback is displayed via audio sounds for all button presses, e.g. ´misfire´ sounds of run scores less than 51, via visual ´bonus conditions´ on the screen and extra points for run scores of 55 and higher (51+1sd). End-of-game feedback is displayed via ´out of ammo´ or ´winner´ animation sequence depending on the participant´s z-scores, being greater than/equal to or less than a preset win threshold.

In the PSI INVADERS each button press actually generates two, 100 bit data samples from the RNG. That is, with every button press a second matched trial is simultaneously sampled, the so-called ´hidden´ sample (or ´noncontingent´ / ´silent´ / ´parallel´ control data sample), of which the participant receives no feedback or information. Identical RNG sampling parameters are employed for both conditions and the order of feedback- and ´hidden´ samples is pseudorandomly determined. Since the ´hidden´ sample is not displayed, special Utility programs, included in the PsiLab package are used to ´read´ it (Apple word processors do the same work).

The setup of present experiment was straightforward. Two principal areas were of interest after considering the disappointing up-to-date results with the supposedly motivating games, the question being if they elicit any PK at all. (i) It was predicted that the American and Icelandic samples would separately and together depart significantly from MCE. In this case, the game was supposed to measure PK ability, which appears either as a hit or miss, and accordingly a two-tailed test was to be used. (ii) The second purpose was to compare the groups of participants in order to look for similarities or differences between them.

METHOD

Experiment One: A Group of subjects from U.S.

This part of the experiment was conducted at the Foundation for Research of the Nature of Man (FRNM), in Durham, North Carolina, where the author was a student at a Summer Study Program in 1985. No exploratory study was conducted since in the game, PSI INVADERS, the computer program is essentially the experimental design.

Subjects

A group of 15 subjects participated in this part of the experiment, mostly volunteers responding to advertisements which were put up in different places at the Duke University Campus in Durham. Of the 15 participants, tested between the 2nd and 20th of July, 5 were males and 10 were females, aged from 9 to 40, the mean age being 26. 4 students in the Summer Study Program at the FRNM, 2 of the staff and 9 outsiders were tested.

Apparatus and setup

The only equipment used were APPLE II microcomputer, connected with a video screen, two disk drives, a steering pedal, the PSI INVADERS disk and data disks. In the computer was placed hardware Bierman-RIPP type electronic noise Random Number Generator board. It converted the analog noise voltage from two independent avalanche electronic noise diodes into two digitized data-bit streams.

Before the test, the experimenter registered the program with prearranged parameters, that served as a design: how many participants were going to play (max participants/study=15), how many games each (max games/participant=3), what ´win threshold´ the player had to overcome in order to win a game (z-score criterion for win=.45), and if RNG hits were to affect the game´s difficulty level (win increment=0). Finally, the experimenter decided the delay between each of the 100 trials/RNG run (delay between RNG trials=5). When the button on the steering pedal was pressed, the gun fired randomly (p=.5), and the proportion of fires and misfires was automatically saved in the program. It was up to the participant how long time each game endured -circa 5 minutes on average.

PROCEDURE

Each participant played three PSI INVADERS games in the experimental room, individually and alone for the latter 'real' games. Only these two games were to be analysed as data, not the first one, which was played as a warm-up exercise for the participant to get used to the computer and possibly lower stress. The participant played the warm-up game in the room with the experimenter, who explained everything necessary and answered questions. The following two games were kept as confirmatory data, played while the experimenter was outside of the room. No additional instructions were made to those appearing on the screen before the game started.

After the first game, the experimenter immediately displayed the results on the screen in graphic form that showed the z-scores distributed above or below chance. After the third game the participant (having played the second and third game in a row) called the experimenter who displayed the z-scores on the graph for the two latter games. When displayed, the participant was first shown the results of the second game and then of the third game.

Experiment Two: A group of subjects from Iceland

The second part of the experiment was conducted in April 18 to May 20, 1986, at the University of Iceland in Reykjavik. This part of the experiment was made identical to the American part in every detail. 15 subjects participated, volunteers responding to advertisements put up in different places on the university grounds, as well as persons that the author personally knew. 8 males and 7 females, of which 7 were students at the University, aged from 19 to 64, the average age being 30, were tested. The instruction, appearing on the screen before the game started, were translated for those who did not understand English.

RESULTS

Since it was never the intention of generalizing the results from the two groups to the U.S. or the Icelandic populations, the participants being volunteers and some of the U.S. group not being Americans by origins, the one-sample t-test for the mean seemed inappropriate. The effect itself is of primary interest rather than the difference between effects. The z-test would probably be used in this case comparing the sample mean to a known theoretical mean.

For the U.S. group the z-test statistic for the mean yielded positive significant results $p<.01$ ($n=30$, $z=2.680$, two-tailed) for the experimental situation of second and third game together. The experimental condition of games two and three together for the Icelandic group did not provide significant results ($n=30$, $z=.164$, two tailed), but the total z-score was in the expected direction. The combined experimental situation of second and third game for both groups together supplied a positive significance of $p<.05$, ($n=60$, $z=2.011$, two tailed). 2 sample t-test analysis for the two groups, i.e. difference in scoring between groups, gave an insignificant pooled variance estimate of $p=.076$ (t value=1.84, df=28, two tailed). Overall z-score among groups are presented in table 1.

Post-hoc analysis on the data showed that all 90 games (game situations 1-3 for both groups) together, provided a probability level of $p<.10$ ($n=90$, $z=1.621$, two-tailed), both conditions together thus yielding marginal positive results.

When one of the participants in the U.S. sample and two of the participants in the Icelandic sample were playing a ´real´ game alone, an error was displayed on the screen accompanied with loud noise. (For description and meaning of this error see the PsiLab manual.) The U.S. participant and one of the Icelandic participants were dismissed from the data pool and other subjects gotten instead. For the second Icelandic participant, being personally known to the experimenter as an honest lawyer, another file was made and he began all over.

The two groups were tested on two different Apple computers with two different RNGs. The PSI INVADERS program was not run in the so-called Simulation (control/matched data) mode, which the PsiLab package can

TABLE 1
The z-scores from PSI INVADERS computer game, for the two groups of subjects. The 2-tailed p value is given for all tests.

	Game 1	Game 2	Game 3	Games 1+2+3	Games 2+3
The group tested in the U.S.	z=.661 n=15 n.s.	z=2.319 n=15 p=.02	z=1.472 n=15 n.s	z=2.226 n=45 p=.03	z=2.680 n=30 p=.007
The group tested in Iceland	z=-.713 n=15 n.s.	z=-.460 n=15 n.s.	z=.692 n=15 n.s.	z=-.240 n=45 n.s.	z=.164 n=30 n.s.
Combined	z=-.052 n=30 n.s.	z=1.315 n=30 n.s.	z=1.530 n=30 n.s.	z=1.621 n=90 n.s.	z=2.011 n=60 p=.044

produce, i.e. separately conducted control trials without a subject present. The internal counterbalancing is therefore only for the first order effects but not for higher order effects of randomness of the RNG. The RNG board used for the Icelandic group was brand-new and the RNG board used for the U.S. group had been tested for a different type of a computer game a few days before the experiment (this was done by using an algorithm that produced digits 1-6). This was a different way of getting random numbers although the RNG was the same (the randomness tests for the U.S. RNG and serial numbers of the RNG boards can be available).

Yet, in the absence of Simulation matched data, it may be argued that the 'hidden' data can be treated as suitable control data. The parallel control trials are indeed subject to the same comparison as the real data. As can be seen in table 2, the run scores from the parallel control trials yielded insignificant results (as one would expect if the RNG was running normally and there was no subject effect in the ´hidden´ data).

TABLE 2
The z-scores from PSI INVADERS computer game

The ´hidden´ parallel control z-scores for the two groups of subjects. Two-tailed p value is given for all tests.

	Game 1	Game 2	Game 3	Games 1+2+3	Games 2+3 experimental situation
The group tested in the U.S.	z=-.930 n=15 n.s.	z=1.456 n=15 n.s.	z=.868 n=15 n.s.	z=.805 n=45 n.s.	z=1.643 n=30 n.s.
The group tested in Iceland	z=.862 n=15 n.s.	z=1.379 n=15 n.s.	z=.238 n=15 n.s.	z=1.431 n=45 n.s.	z=1.143 n=30 n.s.

DISCUSSION

In this experiment participants engaged in an APPLE video computer game, in which an RNG-PK task had been embedded. The U.S. group departed from chance par excellence, the Icelandic group scored slightly in the expected direction and combined the groups yielded significant results. The RNGs apparently behaved as they should do. Thus the purportedly interesting game produced evidence for PK.

What happened with the group tested in Iceland? There was a difference between the two groups though it did not reach a statistical significance (p=.076, t value=1.84, df=28, two tailed). In Iceland computer video games were totally unknown until a few years ago. Only recently they have become popular amongst the younger generation. In the U.S., as far as I know, computer games have been accepted and popular for quite some time. Perhaps this may indirectly indicate that the computer game did not ´appeal´ to the group tested in Iceland.

Computer games have for several years been used to demonstrate micro-PK effects. It is stressed that qualities of the test situation like excitement and ´real´ interest are important for the elicitation of PK. The games are supposed to abolish boredom and motivate the player to do well. It might be pointed out that subjects, showing less interest in playing games would get lower hits than subjects motivated by gaming situations. It would be interesting to test whether computer ´maniacs´ would provide better evidence for micro-PK effect than subjects with no prior experience, the question being if ´maniacs´ become more easily aroused and excited with computergames (which often have some inbuild random processes) than other subjects.

The idea of motivation is old, even in parapsychology. Gardner Murphy (1945, p.82) writes for instance: "In the case of apparitions, too, we have abundant evidence of the role of motivation on the part of the percipient". Unfortunately, it is empirically not possible to treat consequences of a game as an independant variable because of the difficulties in operationalizing concepts like interest and motivation. One can state that different setups/conditions evoke certain effects in type of behaviour. But we can never know if the subject´s interest produces the behaviour. The fundamental question is therefore not if interest can be evoked in an exciting game-like situation, but whether apparently exciting RNG-PK games can produce clearer evidence for PK than other methods and with what subjects.

ABSTRACT

An overview on published RNG-PK computer games is given. Two groups of volunteers were tested in a RNG-PK computer game, called PSI INVADERS; 15 subjects at FRNM in the U.S. and 15 subjects at the University of Iceland. Each participant played one warm-up and two ´real´ games, thus a total of 90 games was obtained. It was predicted that each group, as well as all subjects together, would provide statistical deviations from MCE. Z-score test yielded positive significant probabilities for the U.S. group (p=.007) and for the

combined groups (p=.044). The Icelandic group scored in the expected direction but not significantly. Run scores from parallel control trials yielded insignificant results. Marginal but not significant difference (p=.076) was detected between the two groups.

REFERENCES

Beloff, J., Broughton, R.S., and Wilson,K. (Interim report to the Parapsychology Foundation.), Unpublished, 26 October, 1978.

Berger, R.E., and Honorton, C. ´Psilab//: A stadardized psi-testing system´, In R.A. White, and J. Solfvin (eds.): Research in Parapsychology, 1984, Metuchen, N.J., Scarecrow Press, 1985, 68-71.

Braud, W., and Schroeter, W. ´Psi tests with Algernon, a computer oracle´, in W.G. Roll, J. Beloff and R.A. White (eds.): Research in Parapsychology 1982, Metuchen, N.J., Scarecrow Press, 1983, 163-165.

Broughton, R.S. ´An experiment with the head of jut´, E.J.P., 2, 1979, 337-357.

Broughton, R.S. ´Parapsychology and the experimental method´, in B. Shapin and L. Coly (eds.): Proceedings of an International Conference: Parapsychology and the Experimental Method, Parapsychology Foundation, New York, 1982, 24-35. Also published under the title:´Computer Methodology: Total Control with the Human Face´, Parapsychology Review, 13, 1982, 1-6.

Broughton, R.S. ´The use of computers in psychical research´, in I. Grattan-Guinness (ed.): Psychical Research, Wellingborough, Northamptonshire, Aquarian Press, 1982, 254-264.

Broughton, R.S., and Perlstrom, J. ´A Competitive computer game in PK research: Some preliminary findings´, in R.A. White, and J. Solfvin (eds.): Research in Parapsychology 1984, Metuchen, N.J., Scarecrow Press, 1985a, 74-81.

Broughton, R.S., and Perlstrom, J.R. ´Further studies using a competitive computer PK game´, (Paper presented at the twenty eight Annual Convention of the Parapsychological Association.), Unpublished, 1985b.

Debes, J, and Morris, R. ´Comparison of stiving and nonstriving instructional sets in a PK study´, J.o.P, 46, 1982, 297-312.

Honorton, C. Barker, P. and Sondow, N. ´Feedback and participant-selection parameters in a computer RNG study´, in W. G. Roll, J. Beloff, and R.A. White (eds.): Research in Parapsychology 1982, Metuchen, N.J., Scarecrow Press, 1983, 157-159.

Murphy, G. ´Difficulties confronting the survival hypothesis´, J. A.S.P.R., 39, 1945, 67-94. Also published in:´Three papers on the survival problem´, New York, The American Society for Psychical Research, 1945.

PsiLab, A manual, Psychophysical Research Laboratories (PRL), 301 College Road East, Princeton, N.J. 08540, Copyright 1985.

Ratte, R.J. ´Comparison of game and standard PK testing techniques under competitive and noncompetitive conditions´, J.o.P., 24 ,1960, 235-244.

Ratte, R.J. and Greene, F.M. ´An exploratory investigation of PK in a game situation´, J.o.P., 24, 1960, 159-170.

Rhine, J.B. and Pratt, J.G.: Parapsychology: Frontier Science of the Mind (fifth printing), Springfield, Charles C. Thomas, 1974.

Schechter, E.I., Barker, P, and Varvoglis, M.A. ´Preliminary study with a PK game involving distraction from the psi task´, in W. G. Roll, J. Beloff, and R.A. White (eds.): Research in Parapsychology 1982, Metuchen, N.J., Scarecrow Press, 1983, 152-154.

Schechter, E.I., Barker, P. and Varvoglis, M.P. ´A second study with the "Psi Ball" RNG-PK game´, in R.A. White and R.S. Broughton (eds.): Research in Parapsychology 1983, Metuchen, N.J., Scarecrow Press, 1984, 93-94.

Schechter, E.I., Honorton, C., Barker, P., and Varvoglis, M.P. ´Relationship between participant traits and scores on two computer-controlled RNG-PK games´, in R.A. White and R.S. Broughton (eds.): Research in Parapsychology 1983, Metuchen, N.J., Scarecrow Press, 1984, 32-33.

Steen, D. ´Success with complex targets in a PK baseball game´, J.o.P., 21, 1957, 133-146.

Talbert, R. and Debes, J. ´Time-displacement psychokinetic effects on a random-number generator using varying amount of feedback´, in W.G. Roll, R.L. Morris, and R.A. White (eds.): Research in Parapsychology 1981, Metuchen, N.J., Scarecrow Press, 1982, 58-61.

Varvoglis, M.P. and McCarthy, D. ´Conscious-purposive focus and PK: RNG activity in relation to awareness, task-oriented, and feedback´, J. A.S.P.R., 80, 1986, 1-29.

Weiner, D.H. ´High- vs. Low-risk condition in a computerized PK game´, (Abstract of a paper presented at the fifth Southeaster Regional Parapsychological Association (SERPA) held on February 10 and 11, 1978.), J.o.P., 42, 1978, 64-65.

Weiner, D.H. ´A test of the relationship between PK and motor skill´, (Abstract of a paper presented at the 6th Southeaster Regional Parapsychological Association (SERPA), held on February 9 and 10, 1979.), J.o.P., 43, 1979, 40-41.

Loftur Reimar Gissurarson
University of Edinburgh
Koestler Chair of Parapsychology
7 George Square
Edinburgh EH8 9JZ

EUR.J.PARAPSYCHOL., V.6, 1986

CORRELATED HEMISPHERIC ASYMMETRY IN THE SENSORY AND ESP PROCESSING OF 'EMOTIONAL' AND 'NONEMOTIONAL' STIMULI

Michaeleen Maher
The Graduate School of the
City University of New York

Specialized functioning of the human cerebral hemispheres plays a major role in cognitive activity. This fact has been well documented since Marcel Dax, in 1836, first noted the coincidence between speech defects and lesions in the left half of the brain (Oppenheimer,1977). Although the asymmetric functioning of the hemispheres has been extensively studied, it has been difficult to achieve a full theoretical perspective regarding the contribution each hemisphere makes to ongoing mentation. Problems have arisen from (a) difficulties in integrating the results of data obtained and analyzed by diverse techniques and methodologies; (b) individual and sex differences in the cognitive components of neural organization (Charman, 1979; McGlone, 1980; Witelson, 1976); (c) incomplete specification of parameters such as handedness and familial sinistrality, and their role in the genetic organization of the brain (Hardyck and Petrinovitch, 1977); (d) the effects of motivation, context, and psychological set in modifying or altering hemispheric response proclivities (Gruber and Segalowitz, 1977; Kinsbourne, 1970); (e) a general failure to interpret findings with a view towards the hierarchical ordering of cognitive events (i.e., with sufficient parsimony) (Luria, 1973; Sidtis, 1980); (f) a failure to experimentally address potentially relevant factors such as the activity vs. passivity dimensions of the task, and the discrete vs. continuous nature of the stimuli (Moore and Haynes, 1980); and (g) fundamental limitations inherent in any neuropsychological approach which is constrained to extrapolate the significance of brain events from observable human behaviours.

In spite of the theoretical difficulties, some robust convergence in the data has allowed us to be reasonably certain of a few basic tenets of functional brain asymmetry. Within a model for which the classical prototype is the brain of a male dextral who has no familial history of sinistrality, the portrait which has emerged is that of a left hemisphere specialized for verbal, abstract, analytical processes, as well as those requiring temporal sequencing, as in mathematical calculation, and a right hemisphere specialized for simultaneous gestalt abilities, such as facial recognition, manipulospatiality, depth perception, recogizing rotated forms, performing on visual and tactile mazes, and other forms of holistic processing.

There is also evidence for specialized right hemispheric mediation of human emotional responsivity. The idea that the right hemisphere plays a special role in affective processing was first proposed by Hughlings-Jackson (1879), who noted that language impaired (i.e. left hemisphere) patients continued to express emotion by varying the pitch, tempo and tone of their interjections. More recently Ross and Mesulam (1979) and Ross (1981) have presented clinico-anatomic evidence suggesting that the emotional components of language, including both 'prosody and emotional gesturing', are functionally organized in the right hemisphere. Ross and Mesulam contend that such organization, 'mirrors that of propositional language in the left hemisphere' (Ross, p. 561).

Experimental approaches testing left vs. right field advantages in processing the emotional components of facial recognition (Hansch and Pirozzolo, 1980; Ley and Bryden, 1979; Suberi and McKeever, 1977), as well as studies examining hemispheric advantages in the identification of emotional intonation (Carmon and Nachson, 1973; Haggard and Parkinson, 1971; King and Kimura, 1972) have supported the notion of specialized right hemisphere complicity in emotional processing.

Higher order emotional processing, as well as measured affect have shown themselves responsive to differential hemispheric manipulation. Safer and Leventhal (1977) demonstrated that when participants interpreted verbal passages heard with their left ears, emotional tone cues had greater saliency than did content cues in determining their subsequent evaluations of the passages. Dimond and Farrington (1977) in a novel experiment utilizing a contact lens device (which initially restricted input to the hemifields of a single hemisphere) measured changes in heart rate while showing films to either the left or right hemisphere. A Tom and Jerry cartoon provoked an elevated response when showed to the left hemisphere as opposed to the right, but the film of a surgical operation showed an enhanced response from the right hemisphere. A travel film showed no significant differences in reaction between the two hemispheres. Participants' subsequent rating of the films as humorous, pleasant, unpleasant, or horrific revealed that the Tom and Jerry cartoon and the surgical operation film were both rated as more horrific and unpleasant after being viewed by the right hemisphere. According to the authors, the genuinely enhanced response from the right hemisphere triggered by the unpleasant surgical film suggests that "this kind of emotional experience is more sensitively received" by the right hemisphere and lends "credence to the view that the right hemisphere plays a special role in what is commonly regarded as 'emotion'" (Dimond and Jarrington, 1977, p. 259).

Experimental approaches attempting to use physiological measures as direct indicators of a psi response (i.e., as dependent variables) have successfully used stimuli with an emotional component. Using a plethysmograph, emotional responses to remote stimuli have been detected by Dean (1969); Dean and Nash (1967); Haraldsson (1970); and Schouten (1976). Millar, however, has questioned the validity of most of these findings (Millar, 1979).

Attempts have also been made to use various types of EEG data as detectors of psi experience. In a seminal experiment, Tart (1963) reported evidence of increased EEG complexity in his isolated percipients during times when either Tart himself received painful two second electric shocks, or when the current was passed through a resistor. EEG responses occurring in random two second shock periods were compared to responses occurring in two second control periods sampled eight seconds after each shock had terminated. An electronic period analyzer assessed the complexity levels. Participants, who were asked to press a key when they felt the presence of a shock delivered to the experimenter, were unable to accurately respond to these events on a conscious behavioural level. Although Hearne (1981a) later failed to replicate the remote shock response, his experiment contained important methodological differences (Tart, 1983, pp 221-223)

In another study, average power and peak power recorded from occipital lobes significantly decreased during times when a remote strobe flashed at 16 flashes per second for 10 seconds as compared to periods when it did not flash (Targ and Puthoff, 1974). Five of seven sessions employed an agent (or observer of the light flashes) and all sessions were conducted with a single subject who possessed a 'monochromatic' EEG spectrum. The subject did not exhibit extrachance accuracy when she attempted to consciously discriminate the strobe periods from the control periods. In a replication experiment, which eliminated the agent and was entirely computer controlled, Kelly and Lenz (1976) obtained "encouraging" evidence in support of the alpha blocking effect. However, inconsistencies in subsequent results derived from the same experimental paradigm caused the original authors to lose confidence in their effect, which they eventually conceded was only "suggestive" (May, Targ, and Puthoff, 1979).

Contingent negative variation (CNV) was used by Levin and Kennedy (1975), who attempted to determine whether the CNV depended only upon expectancy, or was a function of the future motor activity to be performed. Participants were expected to press a key as soon as a green light appeared several seconds after a warning signal. If a red light appeared instead, no response was required. A random number generator (RNG) selected which light was to appear immediately before it came on. Significantly more evidence of expectancy appeared in participants' CNV's prior to the green light, but replication attempts failed to confirm the promising results exhibited initially (Levin and Kennedy, 1975; Kennedy, 1979). Hartwell (1978) also failed to detect psi mediated responses in the CNV.

Averaged cortical evoked potential (EP) measures to telepathic stimulation were obtained by Lloyd (1973), who reported that two curves from one subject showed psychic responsiveness similar to an auditory evoked potential. However, when Millar employed essentially the same procedure with 20 subject pairs, and added control conditions (which Lloyd had omitted) he reported that ".. no scrap of evidence was obtained in support of the validity of the Lloyd effect" (Millar, 1976, p. 27).

Hearne (1977) tested whether the degree of emotional closeness between the agent and the subject could affect the EP produced in response to a strobe lamp, when remote agents concurrently viewed photographs of the subjects. In control periods, agents did not view photographs when the light flashed. An analysis of variance revealed a difference in the amplitudes of the earliest of four peaks between experimental and control conditions, as well as a significant interaction between the experimental conditions and the emotional closeness of the eight subject-agent pairs. A later failure to replicate the effect led the author to propose that his earlier success may have been due to statistical anomaly (Hearne, 1981b).

It is clear that an ambiguous pattern of findings has accrued from these pioneering attempts to use sophisticated physiological monitoring techniques to detect unconscious psi responding. This is not very surprising in view of the fact that the replication attempts were rarely methodologically equivalent to the original studies, and experimenter attitudes and motivation may also have differed provocatively.

Some prior work with laterality models of psi functioning (Broughton, 1976; 1977; Maher and Schmeidler, 1977; Maher, Peratsakis and Schmeidler, 1979) can be interpreted as indicating that psi processing utilizes lateralized brain mechanisms in an orthodox manner. That is, psi processing may engage the hemispheres in a manner homologous to the engagement of the hemispheres during sensory processing. It should be noted, however, that Broughton interprets his results differently. He favours a model which hypothesizes a psi-inhibiting mechanism in the left hemisphere which can thwart effective psi performance by the right hemisphere. His experiments have attempted to divert the hypothesized mechanism by engaging the left hemisphere in an irrelevant task. Although the existence of the postulated inhibitory mechanism remains a feasible interpretation of his results, Broughton's studies have not addressed the possibility that potential psi processing in each of the hemispheres may be comparably facilitated by engaging the alternate hemisphere with an appropriate distracting task.

The present experiment was undertaken in an attempt to provide supporting evidence that specialized brain processors are utilized conventionally during psi processing. The study sought to provide evidence for unconscious psi responding, and to compare such responding with the presence or absence of appropriate conscious responses from both participants and blind judges.

Since the abundance of alpha produced by one hemisphere relative to the other during a cognitive task is negatively correlated with that hemisphere's engagement in the task, neuropsychologists have used comparative measures of relative alpha production in both hemispheres during a task to determine which of the hemispheres is predominantly occupied with the processing requirements of the task (Erlichman and Weiner, 1980; Galin and Ornstein, 1972; Mckee, Humphrey and McAdam, 1973; Trotman and Hammond, 1979; Willis, Wheatly and Mitchell, 1979; etc.)

Given stimuli that can elicit hemispheric asychrony during sensory processing, such stimuli should allow us to predict congruous hemispheric asychrony during extrasensory processing. Indeed, the predicted differences in hemispheric activation may themselves be used as evidence for psi.

Since emotion appears to have its own functional coordinates in the right hemispheres of laterized brains, stimuli with emotion inducing properties would seem well suited to the task of evoking greater relative right hemispheric participation during both sensory and extrasensory processing. Several other factors suggest that material with an emotional component might be especially appropriate for use in extrasensory stimulation studies. Emotion has figured heavily as a theme in spontaneous psychic experiences since earliest times. When a systematic study of spontaneous case material was initiated by the Society for Psychical Research in 1889, it was determined that "emotionally charged situations predominated in hallucinatory experiences involving present situations" (Rhine, 1977, p.66). A few studies attempting to assess spontaneous case material with statistical techniques have reported emotional elements present both in descriptions put forth by percipients, and in reactions by sensitives who were brought to the locales and asked to convey their impressions (Maher and Schmeidler, 1975; Moss and Schmeidler, 1968; Schmeidler, 1966).

Given the current right hemisphere findings regarding emotion in cerebral lateralization research, it seems reasonable to expect that emotional elements present in stimulus material may effectively bias the dextral brain towards greater relative right hemispheric activation during psi processing than might nonemotional elements.

Nonemotional material has also proven effective as free response target material in studies of extrasensory perception. Targ and

Puthoff (1977) in a series of remote viewing experiments, have reported substantial success in eliciting psi in the free response reports of participants whose remote agents attended to the architectural design features of public buildings, airports, etc. Because this type of stimulus material, devoid for the most part of any emotional content, has proven effective in stimulating psi awareness, it can be assumed that there is nothing particular about the free-response psi experience which must embody an emotional component. In deference to the demonstrated ability of such material to evoke corresponding mentation in the conscious minds of remote participants, architectural themes can be expected to be well suited to the task of inducing nonemotional mentation extrasensorially.

The present study hypothesized that greater relative right hemispheric participation in processing would be exhibited in the EEG data of dextrals who viewed 'emotional' videotaped material than would be exhibited when the same participants viewed 'nonemotional' material, and that corresponding differences in hemispheric activation would prevail when participants attempted to respond to the same videotaped material clairvoyantly.

However, because the cortical processing mechanisms of females are presumed to be either less lateralized or differently lateralized from those of males (see McGlone, 1980, for an excellent review and summary of the evidence) females in the present study were expected to exhibit the predicted differences in processing to a significantly lesser extent than would males.

It should be understood that, with respect to group differences in the processing of the 'emotional' and 'nonemotional' videotaped material, the present study was exploratory in its sensory capacity, as well as in its extrasensory capacity. Prior research had not, after all, demonstrated how normal dextrals would process the two videotapes employed as stimuli. It was recognized from the outset, however, that should the 'emotional' and 'nonemotional' stimuli fail to elicit the predicted group differences in hemispheric processing, data would none the less have been collected which would enable the primary hypothesis of congruence in the lateralized processing of sensory and ESP information to be tested with individuals' patterns of hemispheric activity.

In addition to the anticipated evidence for psi processing from participants' unconscious patterns of brain activation, significant evidence of psi was expected to be exhibited in the blind judging analysis of participants' conscious mentation reports. The reports were expected to be matched to their respective target videotapes more successfully than chance would allow.

METHOD

Participants and experimenters

Participants were 20 adult dextral volunteers, who espoused an open, accepting attitude towards the prospect of ESP success. Ages ranged between 18 and 45. There were 10 males and 10 females. Two experimenters, one female (M.M.) and one male (K.R.) recruited participants, administered tests, and monitored the laboratory equipment.

Design

A within-subject design was employed, with each participant having four task conditions. In the two primary ESP conditions, participants attempted to use clairvoyance to fathom the contents of two videotapes, which were played in random order in a distant room. One tape contained material with emotion inducing properties, and was expected to bias the activation of participants' cerebral hemispheres towards relatively greater right hemispheric participation in processing than would the other videotape, which contained material generally devoid of emotion inducing content. EEG data from three pilot participants (two males and one female) had indicated that the hemispheric activation propensities of the two videotapes were in the requisite direction when dextrals viewed the tapes on a TV screen. Because the EEG facility was available for only a relatively limited period of time, the running of a more substantial pilot series was precluded.

Integrated EEG activity was monitored from temporal/parietal sites (referenced to a common vertex) in both left and right hemispheres during two extrasensory viewing periods. The electrode placement was the standard 10/20 procedure used in the laboratory, which specializes in the study of lateralized hemispheric function. After each ESP viewing period, participants gave mentation reports which were tape recorded for later analysis by outside blind judges. In two subsequent sensory conditions, EEG data were similarly recorded while participants viewed the two randomly sequenced videotapes on a TV monitor. Participants' electrodes were not removed during the interval between the ESP and the sensory sessions.

To check on participants' laterality profiles, additional EEG data were collected while participants performed alternately sequenced drawing and writing tests for about two minutes each. These tests, which have shown themselves to be robust with respect to indexing lateralized functioning (Davidson and Ehrlichman, 1980) were standard operating procedure in the laboratory where the experiment was conducted. Indices of integrated alpha production (amplitude x time) filtered from participants' EEG activity, allowed for comparisons to be made between relative hemispheric engagement while occupied in the two laterality tasks, and also while viewing the two different videotapes in the sensory and extrasensory conditions.

Equipment

The experiment was performed in the EEG laboratory located on the sixth floor of the Graduate Centre of the City University of New York. The laboratory houses an Industrial acoustics chamber (series 1200) which is 8 ft x 15 ft, with 1 ft thick walls and electromagnetic shielding. Standard intercom and cassette tape recording equipment for recording the speech of participants was employed. In addition, the chamber contains a comfortable reclining chair, which was utilized for both the sensory and extrasensory perception of the target material. A 3/4 in. videocassette deck and monitor were installed in the chamber for use in the experiment. Beckman silver-silver chloride disk electrodes were affixed with Grass electrode cream to bilateral temporal/parietal sites (midway between T3 and P3; and T4 and P4) referenced to a common vertex (Cz). Electrode resistances were always less than 9K Ohm. A ground electrode was placed in the centre of the forehead.

Outside the chamber, the laboratory contains a Beckman Type R polygraph and a Vetter 8-channel Model C8 FM cassette tape recorder for recording EEG directly onto cassette tapes. For this experiment, integrated alpha production was averaged over 15 second intervals, providing a series of approximately 60 raw scores for each hemisphere in each of the four conditions of the study. A separate room, located beyond the stairwell adjacent to the laboratory, was equipped with an additional 3/4 in. videocassette deck and monitor for presentation of the videotaped target material during the extrasensory conditions.

Materials

Four videocassettes were prepared -two for use in the sensory conditions, and two for use in the extrasensory conditions. Each hour-length cassette contained a pairing of the emotional and nonemotional videotaped material. A ten-minute interval of blank tape separated the two stimulus presentations, each of which was slightly less than 20 minutes in length. A one-minute interval of blank tape was contained at the head of each cassette. The ESP conditions utilized two cassettes (whose jackets were labeled ESP-A and ESP-B). One cassette contained the emotional material first and the nonemotional material second. The other cassette contained the nonemotional material first and the emotional material second. The sensory conditions similarly utilized two cassettes (whose jackets were labeled Sensory-A and Sensory-B), each containing a pairing of the emotional and nonemotional material. As with the ESP cassettes, one of the sensory cassettes contained the emotional material first and the nonemotional material second, while the other cassette contained the nonemotional material first and the emotional material second.

A fifth hour-length cassette was prepared as a control tape for participants to view during the ESP conditions. This tape contained no image, but allowed the lit TV screen to be used as a focussing device, and encouraged participants to maintain a proper head and body orientation, and to keep their eyes opened during the ESP viewing periods.

Participants' extraversion was evaluated with a shortened version of the Eysenck Personality Questionaire (EPQ). The short form of the Annette Handedness Questionaire assessed participants' handedness preferences for 12 skills, and inquired into familial characteristics of handedness. Participants' eye dominance (operationalized as eye preference when sighting through a portable telescope) and thumb dominance (operationalized as left or right thumb supremacy when interlocking fingers with clasped hands) were ascertained by the experimenters. (note 1).

A relaxation procedure, modified from Masters and Houston (1972, p. 8-9) was administered to participants prior to each ESP viewing period.

Stimuli

Both videotaped presentations were the conceptions of artists, who graciously lent tapes to the author for use in the experiment. The tape which was expected to induce relatively greater right hemispheric involvement during processing contains a series of interviews with seven males and females who, often falteringly but sometimes eloquently, discuss their varying attitudes towards love and loving. The extemporaneous dissertations are generally sad, intimate, and emotionally evocative. Introspective reports from viewers indicate that the material incites empathy, is emotionally compelling, and is often profoundly depressing. The interviews take place in real time, and the tape contains no within-sequence editing. The accompanying music track carries a ballad -I'm in the Mood for Love- which weaves plaintively throughout the interviews. The tape (designated the Love Tape for experimental purposes) is monochromatic, or black and white.

The alternate color tape, which was expected to induce relatively less right hemispheric engagement during processing, is primarily intellectual in its appeal, and features an analytical approach to its subject matter, New York City, which is depicted as a giant, automated organism. This presentation (designated the City Tape) contains numerous aerial and landscape views of the metropolis, with its pulsing arteries of traffic and distant figures scurrying out of subway exits. Such macrocosmic scenes are contrasted with highly magnified microcosmic views, such as red blood cells moving, as if purposefully, through capillaries. The tape contains many sequences where fast-motion cinematography is utilized to make visual statements central to the artist's thesis. Enormous architectural structures for example, are demolished and re-erected in the blink of a time-elapsed eye. The tape spans an imaginary daily cycle in the life of the city organism. The material is highly edited; camera angles are constantly shifting and changing; and the accompanying sound track is electronic or voice automated.

In the search for appropriate free response stimulus sets to be used alternately as target material in an ESP experiment, perhaps the primary condition that must be met is that the targets must be as dissimilar as possible. In this respect, the two stimulus videotapes seemed ideal. The two tapes could reasonably be distinguished in terms of at least 10 conceptual categories. These were: 1) black and white vs. color; 2) intimate vs. remote, aloof; 3) animate vs. inanimate; 4) real time vs. altered time; 5) close up faces vs. distant figures; 6) inside vs. outside; 7) concrete vs. abstract, analytical; 8) feeling vs. thinking; 9) soft round shapes vs. hard angular shapes; 10) viewers' perspective static vs. viewers' perspective shifting.

Although these conceptual dichotomizations inevitably included some overlap, with respect to basic dissimilarity the limitations of the tapes seemed tolerable. A more problematical requirement of the stimulus sets for this experiment, was that they must also appeal to different specialized processors in the hemispheres. It was recognized that any videotape would necessarily include a substantial visuospatial component. When watching TV one is, after all, looking at a screen in which images are moving around. Theoretically, then, the use of videotapes as targets meant that a built-in bias could be expected to favor relative right hemispheric processing. Given that cognitive processing might be relatively lateralized to the right hemisphere, the question became -would the negative emotional approach of the Love Tape thesis engage the right hemisphere relatively more (with respect to the left) than would the abstract, intellectual approach of the City Tape thesis? It was 'relative' right hemispheric differences in processing that were important here, rather than any right vs. left hemisphere differences per se. Both tapes contained speech and both contained music. In the Love Tape, the speech was faltering, prosodic and emotionally meaningful, while the music was a poignant ballad. The City Tape, in contrast, contained automated speech which was crisp, terse and robot-like, while the music was atonal and electronic.

A further factor in the Love Tape which might be expected to load on the right hemisphere was the fact that, throughout the tape, the viewer watched faces. Facial processing (i.e., the brain's ability to holistically document facial information) was among the first of the explicit specialized functions to be attributed to the right hemisphere.

Therefore, despite the theoretically equivocal nature of any form of continuous, multiplex processing, it was hypothesized that the sum of the parts of the Love Tape stimulus would bias dextrals' right hemispheres towards relatively greater right hemispheric involvement in processing than would the sum of the parts of the City Tape stimulus.

In any event, should the selected videotapes fail to elicit the predicted group differences in lateralized processing, the principal hypothesis of congruence in lateralized activation during sensory and extrasensory processing of the same stimuli could still be tested with individuals' patterns of activity. In this eventuality, the wide variability which could be anticipated for individuals' processing of the videotapes would become an asset rather than a disadvantage.

Procedure

The experimenters met each participant at the laboratory at the appointed hour. After a brief tour of the facility, including the acoustics chamber, and the room where the ESP stimulus tapes were to be played, the participant filled out the handedness and extraversion questionaires.

After being settled comfortably in the acoustics chamber, the participant had electrodes affixed, usually by both experimenters working simultaneously. Four electrodes were employed -one on the temporal-parietal midpoint of the left hemisphere; one on the temporal-parietal midpoint of the right hemisphere; one central vertex reference electrode and a ground electrode in the centre of the forehead.

After relaxing for a brief period, while polygraph checks were made, the participant engaged for about two minutes either in writing an intellectual letter to a friend, or in drawing the picture of an imaginary person. Subsequently the alternate task was performed. EEG data were collected while the participant was occupied in the tasks.

A brief rest period followed, after which the ESP control cassette was inserted into the video deck, and the monitor turned on. The screen lit up, and the background video white noise was adjusted to a comfortable level. The ambient room light was turned off, and the experimenters exited, sealing off the chamber. Tape recorded instructions were played for the participant over the two-way intercom. She or he was thanked for participating in the experiment; given general instructions on mobility and eye fixation; and encouraged to feel that ESP success was indeed possible. A relaxation procedure followed, and after a final countdown, the participant was instructed to "... begin viewing now".

Meanwhile, while M.M. monitored the polygraph equipment, K.R. at an explicit cue during the relaxation instructions, went to the ESP projection room down the hall. The toss of a coin determined which of the two ESP cassettes was to be taken from the drawer and placed in the video deck. (Prior to each participant's arrival, M.M. had used a coin toss to determine whether or not to switch the jacket labels on the two ESP cassettes. The cassettes were identical, but their jackets were labeled ESP-A and ESP-B. This procedure kept K.R. blind as to the order of the stimulus presentations within each cassette.) After putting the selected cassette in the deck, and turning the monitor on, K.R. started a pocket stop watch, and returned to the laboratory before the relaxation instructions had ended. The one-minute blank lead-in at the head of each cassette insured that he remained blind as to the order of the videotaped presentations for each participant.

K.R. performed the duties of time-keeper during the experiment, and advised M.M. when the first stimulus presentation had begun, whereupon M.M. began recording the participants' EEG. K.R. also operated the intercom system, switching it from two-way to one-way during the ESP viewing periods. When K.R. gave the signal that the first videotaped presentation was over, the EEG recording was turned off, and the two-way intercom turned back on.

The participant was then asked to report his or her mentation during the ESP viewing period. Reports varied in length from 2 1/2 minutes to about 7 1/2 minutes, and were tape recorded for later analysis by outside blind judges. When the participant, after several promptings, could remember nothing additional about what had been going through his or her mind during the viewing period, M.M. gave extemporaneous relaxation instructions (the length of which varied inversely according to the length of the mentation report) and a countdown for the second ESP viewing period began. K.R. as timekeeper, again gave the signal marking the start of the second videotaped presentation, and switched off the two-way intercom.

When the second tape was over, the participant provided a second mentation report, and then was allowed a rest period of approximately 10 minutes. The acoustics chamber door was opened, and the room light turned on. The participant was offered a glass of water; smoked if she or he chose, and/or chatted with the experimenters.

Because it was deemed necessary for participants to view the forthcoming sensory videotapes without comparing them to their prior mentation reports, a mild ruse was adopted. The two sensory cassettes (labeled Sensory-A and Sensory-B) were procured from a table behind the participant. Experimenter M.M. explained that the next phase of the experiment involved trying to see how the participant's brain normally behaved when watching TV. It was suggested that each of the two videocassettes held "... a couple of different videotapes", and that it really mattered very little which cassette was used. Therefore the participant was invited to select either cassette for his or her own "... viewing pleasure." The participant was thus left with the impression (although it was not explicitly stated), that the two cassettes contained four different videotapes which were unrelated to the ESP videotapes.

The control cassette was then removed from the deck, and the participant-selected cassette was inserted. When turned on, the one-minute blank lead-in at the head of the tape provided adequate time for M.M. to exit from the chamber and to turn on the EEG recorder as the first sensory presentation began. The intercom was turned to one-way, and the EEG data were recorded during each of the two sensory viewing periods, which were separated by a 10 minute rest interval. During this interval, participants again smoked if they wished, and conversed with the experimenters. Eye dominance and thumb dominance were tested generally by K.R. during the rest interval.

After the second tape was over, the polygraph, tape recorder, and intercom were turned off and M.M. entered the acoustics chamber, sealing the door behind her. After apologizing for the ruse, M.M. explained that the two tapes the participant had just viewed were in fact the same tapes that were played during the two ESP viewing periods, but that the order in which the two ESP tapes had been played was, as yet, unknown to anyone. Participants then reviewed their mentation reports and made a decision about which tape they believed had been playing during each of the ESP viewing periods. M.M. recorded the participant's decision, and then opened the door of the acoustics chamber.

Meanwhile, K.R. went down the hall to the ESP projection room and determined, by rewinding the tape, the order in which the tapes had been played. When he returned, he announced the tape order, which was recorded. The participant was debriefed about the explicit nature of the experiment; electrodes were removed; and he or she exited from the laboratory.

EEG Analyses

Each participant's EEG activity was collected on one side only of two 120 minute portable cassette (60 minutes per side). One cassette held the EEG from the laterality profile tests and the two ESP mentation periods, while the other cassette held the EEG from the sensory viewings of the two videotapes. A 50 uv 10 Hz oscillator generator sine wave signal was recorded for approximately two minutes at the head of each participant's initial cassette, for use as a reference signal. The reference signal was placed on the tape after the FM recording channels had been calibrated so that the oscillator signal registered a meter reading of zero decibels on each of the recording channels used to collect the EEG data.

The conversion of the taped EEG to a digitalized output was accomplished with a custom built filtering and integration EEG analyzer. Each of the two integration channels (one of which was to process EEG from the left hemisphere and one of which was to process EEG from the right hemisphere) was set so that the reference oscillator signal summed to 650 arbitrary units for every 15 seconds of recording. Raw EEG from the left and right hemispheres was then filtered for 8-13 Hz (alpha) activity and an averaged integration measure (amplitude x time) of the rectified waveform was digitally displayed for each hemisphere with reference to the same arbitrary reference signal.

Artifacts in the EEG recording were identified through a blind inspection of the polygraph records, and any epoch containing an artifact was removed from the digitalized output by means of the correponding footage numbers. Records were relatively artifact free. Minor variations in lengths and onsets of the tapes on the different cassettes, as well as the need to exclude the titles at the beginning of each tape required that four additional epochs (one minute) from the head and tail of each data set be excluded from the analysis.

Blind Judging

To study the process of judging extrasensory material, a three-fold procedure was employed with two experienced blind judges. The mentation report recordings (collected on portable cassettes) were transcribed and randomized. Judges were asked initially to compare each of the mentation reports (a total of 40 reports were presented) with both the Love Tape and the City Tape, and to make a global determination about which tape had been playing during the ESP period corresponding to each mentation report. In this first stage of the judging process, judges believed that there was an equal probability that either tape had been played for any given report (N=40).

After the scoring had been collected, judges were given participants' paired mentation transcriptions, in the order of their having been rendered, and asked to make a second determination about which tape had been played during the corresponding ESP periods (N=20). This process invariably necessitated some change in the judges' former ratings (which they were asked to ignore) because now the judges understood that both tapes had been played for each of the 20 participants -one during the primary ESP period, and the other during the secondary ESP period.

After the second batch of scores were collected, judges were given a list of the 10 bipolar descriptor continua with instructions to use them as categorical aids in studying the mentation reports. After rating the percentage of correspondence of each report along every continuum, and scrutinizing the overall configuration of each report's ratings, judges made a third or final decision about which tape they believed had been playing during each of the participant's ESP periods.

RESULTS

Blind Judging

Results for participants' blind matching of their own mentation reports to the videotapes were at chance, with 12 of the 20 participants (six males and six females) correctly matching their reports to the tapes. Results for the two outside blind judges are presented in Table 1.

TABLE 1
Proportions of ESP mentation reports correctly matched to targets by two judges in three successive trials*

	Trial 1	Trial 2	Trial 3
Judge A	.525	.65	.80
Judge B	.475	.65	.65
	(N=40)	(N=20)	(N=20)

* MCE=.5

It can be seen from the table that one of the judges improved in accuracy over the three-fold process of judging, while the other judge improved from the first to the second step of the procedure, but maintained the same level of accuracy (although individual judgments changed) for the final judging. A z-score analysis evaluating the number of correctly matched reports for Judge A's final judging yields a $z=2.46$, $p<.007$, one-tailed (corrected for continuity). Judge B's final judging, while yielding an excess of appropriate judgements, was not significant.

To determine if there were significantly more correctly matched than incorrectly matched mentations when participants' and judges' final scoring was combined, the 10 tandem reports which individual participants and both judges had correctly matched were contrasted with the three reports which participants and both judges had incorrectly matched. The binomial expansion gives $p<.05$, one-tailed (corrected for continuity), a difference which favors the hypothesis of psi.

A chi-square analysis was also performed to evaluate the expected and obtained frequencies of hits for the three scorers' final judgments. These data are presented in Table 2.

TABLE 2
Binary hits for three scorers in a final evaluation of 20 ESP mentation reports.

Number of hits from three scorers	Expected frequency	Obtained frequency
0	2.5	3
1	7.5	3
2	7.5	4
3	2.5	10

Chi-square gives 26.93, df=3; $p<.001$, showing that multiple judgments of the same protocols tended towards the same conclusion. If each of the twenty summed final scores from three scorers is considered as a unit, with the known probability of a hit being 1.5, a z-score analysis gives $z=1.98$; $p<.03$, one-tailed, confirming that psi is reflected in the three scorers´ final judgments. Table 3 contains participants´ order and final judging data.

To meet the concern that an overall assessment of psi in the blind judging scores ought to be based on an evalution of all of the judgments rather than just the final ones, judges´ seven sets of judgments (three sets of judgments from each of two outside judges and one set of judgments from 20 experimental participants) were pooled and evaluated with a one-sample t-test. Pooled scores for the analysis were derived as follows. For the judges´ first series of 40 judgments, where a participants´ two mentations may have been matched to the same videotape (causing one match to be a hit and the other to be a miss) a score of .5 was accorded to each correct match. For trial 1, then, a judge´s score could range from zero (neither mentation correctly matched) to .5 (one mentation correctly matched) to one (both mentations correctly matched.) For the second and third judging trials, each of the judges´ successful tandem matches (N=20) was accorded a score of one. Unsuccessful matches received a score of zero. Participants (whose judging task was equivalent to the judges´

TABLE 3
Order of testing and final judging results for 10 males and 10 females.

Male Participants

Participant	1	2	3	4	5	6	7	8	9	10
Order of testing	2	3	6	8	10	11	15	16	18	20
Draw/Write order	DW	DW	WD	WD	WD	DW	DW	WD	DW	WD
ESP order	LC	LC	CL	CL	CL	CL	CL	LC	CL	LC
Sensory order	LC	CL	LC	CL	CL	LC	LC	CL	LC	CL
ESP target assignments	LC	CL	CL	CL	LC	LC	CL	CL	CL	LC
Final judging										
Judge A	H	M	H	H	H	M	H	M	H	H
Judge B	M	M	H	H	M	H	H	M	H	H
Participant	H	M	H	H	M	M	H	M	H	H

KEY:
D=Draw a person L=Love Tape H=Hit
W=Write a letter C=City Tape M=Miss

Female Participants

Participant	11	12	13	14	15	16	17	18	19	20
Order of testing	1	4	5	7	9	12	13	17	19	14
Draw/Write order	DW	DW	WD	DW	DW	WD	DW	DW	WD	WD
ESP order	LC	LC	LC	LC	LC	CL	LC	CL	CL	LC
Sensory order	CL	LC	CL	CL	CL	CL	CL	CL	LC	LC
ESP target assignments	LC	CL	CL	LC	CL	CL	LC	LC	CL	LC
Final judging										
Judge A	H	H	H	H	H	H	H	M	H	H
Judge B	H	H	H	H	M	H	H	M	M	H
Participant	H	M	M	H	M	H	H	M	H	H

KEY:
D=Draw a person L=Love Tape H=Hit
W=Write a letter C=City Tape M=Miss

second trial) also received a score of one for correct matches and zero for incorrect matches. Each participants' score was pooled with the six scores from the two judges. The resulting set of 20 scores was subjected to a one-sample t-test (with MCE=3.5). The result, $t=1.74$; $df=19$; $p<.05$, one-tailed, gives supporting evidence that psi was exhibited in the blind judging comparisons.

Missing EEG data

Noise factors rendered one female participant's EEG recording during the sensory viewing of the City Tape unusable. EEG information from only five of a possible 66 epochs could be redeemed, which sampled only the first 1 1/4 minutes of viewing. Consequently, EEG data from her four experimental conditions, as well as EEG data from her laterality profile tests were excluded from the analysis.

EEG Laterality Profile Tests

An analysis of variance (ANOVA) contrasting hemispheric activation differences for the 19 remaining participants while drawing a person and writing a letter showed that participants, as a group, were appropriately lateralized for these tasks (hemisphere x task interaction significant at $F(1,17)=12.34$; $p<.003$). Participants' mean alpha raw scores are presented in Table 4.

It can be seen from the table that participants produced less mean integrated alpha in the right hemisphere while drawing than they did in the left hemisphere, indicating their relatively greater right hemispheric involvement in the drawing task. Conversely, participants while writing exhibited less alpha in the left hemisphere than they did in the right hemisphere, indicating the left hemispheric activating propensities of the writing task. Females showed less of the predicted lateralization effect for these tasks than did males. An ANOVA contrasting hemispheric differences for the 10 males showed the anticipated task x hemisphere interaction significant at $F(1,9)=15.75$; $p=.004$, while the equivalent task x hemisphere interaction for the nine females was only suggestive at $F(1,8)=4.66$; $p<.07$.

TABLE 4
Mean left hemisphere (LH) and right hemisphere (RH)
Alpha raw scores on two laterality tests (N=19)

	LH	RH
Draw	63.855	58.656
Write	57.357	63.749

EEG ANOVA Analyses

A repeated measure ANOVA on participants' mean indices of left hemisphere and right hemisphere alpha production across the sensory and extrasensory tasks revealed that substantially more alpha was generated during the ESP condition as compared to the sensory condition, $F(1,17)=8.41$; $p<.01$. The ANOVA further determined that for the complete body of data, males' hemispheres reacted differently to the videotapes than did females' (sex x tape x hemisphere interaction significant at $F(1,17)=4.62$; $p<.047$). Group means are presented in Table 5.

Separate ANOVAs on males' data alone demonstrated that although males showed the predicted effect of greater relative right hemispheric processing for the Love Tape that for the City Tape in the Sensory condition at a marginal level of significance ($F(1,9)=4.16$; $p<.059$), their ESP alpha did not show an independently significant tape x hemisphere interaction. Nor was the repeated measures (i.e., sensory and ESP data) tape x hemisphere interaction independently significant for males. Females data showed relatively greater differences in alpha between the left and the right hemispheres when processing the City Tape than when processing the Love Tape in both the sensory and ESP conditions. For females, then, the City Tape tended to be relatively more 'right hemispheric' than did the Love Tape. However the effect of greater right hemisphericity for the City Tape was only suggestive across the sensory and ESP data: $F(1,)=3.78$;

TABLE 5
Males′ and females′ mean left hemisphere (LH) and right hemisphere (RH) aplha for the sensory and ESP sessions of the Love Tape and the City Tape

	Sensory			ESP		
	Love	City	Diff	Love	City	Diff
Males						
LH	92.20	83.09	9.11	102.35	96.98	5.37
RH	85.10	78.97	6.13	97.12	94.03	3.09
DIFF	7.10	4.12		5.23	2.95	
Females						
LH	127.37	125.70	1.67	143.65	152.24	-9.59
RH	121.53	116.45	5.08	139.39	142.91	-3.52
DIFF	5.84	9.25		4.26	9.33	

$p<.09$. Neither sensory nor ESP data for females showed a significant tape x hemisphere interaction when considered alone.

When males′ and females′ sensory data is considered together, the tape x hemisphere x sex interaction is suggestive at $F(1,17)=3.18$; $p<.10$ in contrast to the ESP interaction of $F(1,17)=1.75$; $p<.20$.

EEG Correlational Analyses

To determine if the hypothesis of congruence in the lateralized processing of sensory and ESP information would be confirmed by individuals′ laterality patterns (if not by robust group differences in the lateralized processing of the two stimulus videotapes), a correlational analysis was performed. For this analysis, laterality scores were derived for each participant according to the established protocol at the laboratory where the experiment was conducted. Laterality means were derived for each participant from laterality

indices computed for every 15 second interval of viewing in the two sensory (i.e., Sensory Love; Sensory City) and two extrasensory (i.e., ESP Love; ESP City) conditions. The amount of alpha produced by the right hemisphere was subtracted from the amount of alpha produced by the left hemisphere and the resulting difference was divided by the amount of alpha produced by both hemispheres during the same interval. Left hemisphere and right hemisphere alpha raw scores could thus be collapsed into a single laterality index for every 15 second epoch of each session. Treatment means were then derived for each participant in each of the four treatment sessions. With this method, a relatively higher mean was indicative of greater relative right hemispheric participation in processing. For example, if an individual showed a higher laterality mean for Sensory Love than she or he had shown for Sensory City, this indicated that the individual had displayed greater relative right hemispheric participation in processing the Love Tape as compared to the City Tape. Participants' mean laterality scores for the Sensory and ESP sessions of the Love Tape and the City Tape are presented in Table 6.
Males' and females' group means are presented in Figure 1.

In order to correlate individuals' hemispheric activity patterns in the sensory and ESP conditions, each participant's mean laterality index for the City Tape was subtracted from his or her mean laterality index for the Love Tape in the Sensory condition, and again in the ESP condition. The resulting difference scores were then correlated for the 19 participants. The Pearson correlation coefficient of the difference was $r=.4327$; $p<.04$, one-tailed (note 2).

Separate Pearson product moment correlation coefficients were then derived for (a) the group of participants whose ESP mentation reports had been correctly matched to the videotapes by all three judges, and (b) all other participants. It may be recalled that 10 of the 20 participants gave tandem mentation reports which were correctly matched to the videotapes by each of the respective participants and by both outside judges, in the final judging analysis. However, one of the participants whose mentation reports received such correct consensus judging was the female whose inadequate EEG data had been eliminated from the analysis. For the nine remaining participants in the sample of 'best' participants then, the correlation coefficient was $r=.7149$; $p<.015$, one-tailed. The data of the 10 participants whose reports had received incorrect matches from at least one of the three judges yielded a correlation coefficient of $r=-.1348$ (n.s.) (note 3). The difference between the two coefficients was significant ($z=1.86$;

TABLE 6
Mean laterality indices for males' (1-10) and females' (11-19) sensory and ESP sessions of the Love Tape and the City Tape*

Sbj	Sensory Love	Sensory City	Sensory Dif	ESP Love	ESP City	ESP Dif
1	-.019	-.018	-.001	-.044	-.076	+.032
2	+.038	+.025	+.013	+.015	-.009	+.024
3	+.169	+.096	+.073	+.148	+.115	+.033
4	+.131	+.127	+.004	+.066	+.055	+.011
5	+.073	+.080	-.007	+.084	+.067	+.017
6	+.110	+.156	-.046	+.140	+.130	+.010
7	+.043	+.027	+.016	-.029	-.018	-.011
8	+.013	-.009	+.022	-.032	-.064	+.032
9	-.054	-.038	-.016	-.055	+.038	-.093
10	+.031	-.016	+.047	+.100	+.045	+.055
11	+.089	+.205	-.116	+.106	+.173	-.067
12	+.024	+.003	+.021	+.108	+.199	-.091
13	+.138	+.170	-.032	+.127	+.142	-.015
14	+.086	+.093	-.007	+.008	+.047	-.039
15	+.000	+.031	-.031	-.051	-.043	-.008
16	-.017	-.011	-.006	-.106	-.081	+.002
17	+.246	+.125	+.121	+.176	+.154	+.022
18	+.057	+.056	+.001	+.077	+.067	+.010
19	-.058	-.028	-.030	-.004	-.024	+.020

Note: Data were rounded to three decimal places although seven decimal places were used for computation.

$p < .05$, one-tailed.).

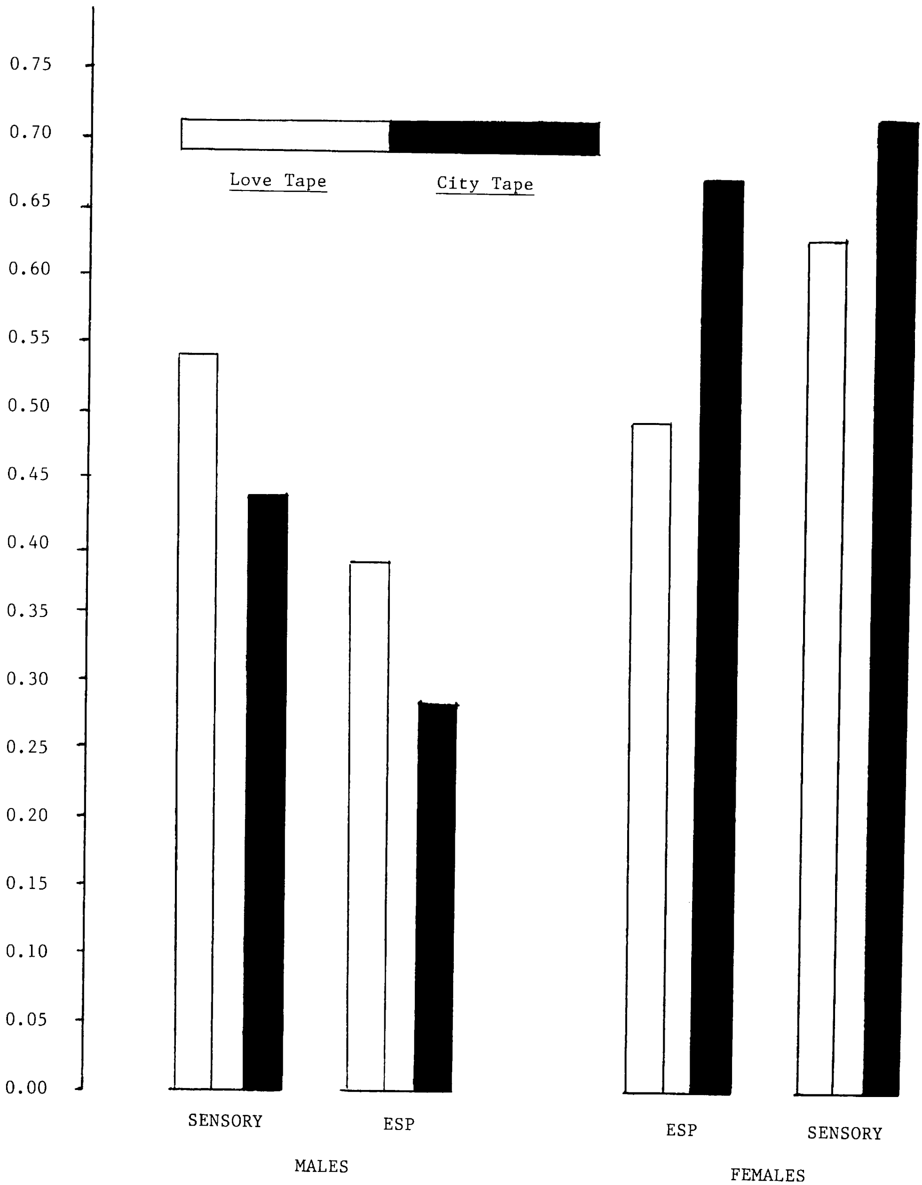
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Love Tape
City Tape
SENSORY
ESP
MALES
ESP
SENSORY
FEMALES

FIGURE 1

DISCUSSION

Blind judging

The chance results for participants -attempting to match their own ESP mentations to videotape targets- lend no support to the ESP hypothesis. It is unfortunate that participants were not able to engage in a variation of the outside judges' three-fold judging procedure, to see if their judging might have improved with repeated tries. However such a procedure (which was never considered for the participants) would have been impractical because the participants were invariably fatigued after all of the data had been collected. It would have been inconsiderate (and perhaps even inhumane) to have subjected them, at this juncture, to a demanding and time-consuming judging process.

Using this process, an outside judge achieved scores which were consistent with the hypothesis of psi. While it has been observed that the true probability of judge A's final matches occurring by chance has not been adequately addressed by comparing only her final judgments to chance expectation, if we were to triple her final probability value to accomodate the hypothesized chance fluctuations of two prior sets of 'guesses', the result would still meet the criterion for significance. If we in fact were to septuple the value to account for all seven sets of presumed guesses (three sets from each judge and one set of participants' guesses), the result would continue to meet the criterion for significance. Moreover it should be understood that while judging instructions encouraged flexibility in making target/protocol assignments, the judging process ultimately led to a set of final irrevocable judgments. These final judgments necessarily reflected and were contingent upon all of the mental operations which had preceeded them.

The improvement in judging during the three-fold process of evaluating participants' mentation is a curious trend. Although judging was nonsignificant for the initial global matchings of the 40 mentation reports to the videotapes, both judges improved in accuracy when evaluating participants' tandem mentations. After working through the second judging procedure, Judge A anecdotally remarked that it seemed to her as if some of the percipients were ´Love Tape types´ while others were ´City tape types´. Judging from the improvement in scores, evaluating percipients´ mentations in tandem appears to have allowed for a finer discrimination of the extrasensorially relevant aspects of the reports.

Judge A also expressed enthusiasm for the third aspect of the judging procedure, which required judges to analytically review the mentations and quantify each with respect to the percent presence or absence of numerous descriptive characteristics. Prior to receiving feedback for her judging (and before any of the data had been analyzed) judge A declared that she was excited about the categorical procedure, and felt that it had helped her to evaluate reports which she had previously considered problematical. Judge B, on the other hand, found the analytical procedure "tedious" and "time consuming". She reported that she had found it difficult to apply some of the scales to the reports, and that she hadn´t felt that the procedure -which required her to "weight things less intuitively"- would produce the best results. Judge A´s scoring improved after using the analytical procedure, while Judge B´s did not. Whether judge A´s improvement in judging accrued as a result of the procedure, or simply as a function of the additional time spent mulling over the mentation reports remains for future research to clarify. It is perhaps significant that Judge A´s timesheet indicated that she had spent a total of 20 hours assessing the mentation material, as compared with a total of 14 hours spent on the project by Judge B.

Judges´ introspective comments (written during the blind judging process) indicated that there was variation in what each judge considered to be salient in the percipients´ reports. Also in some cases judges appeared to be swayed in favor of a correct matching by only one or two isolated comments in the report. Judges´ strategies for matching the mentation reports to the videotapes are the subject of continuing study -and a replication of the judging component of the present experiment has been undertaken.

EEG Laterality Profile Tests

The results of the drawing and writing tests (with males´ hemispheres showing more right hemispheric involvement in the drawing as opposed to the writing task, and females´ hemispheres showing the same effect to a lesser extent) indicate that the sample of dextral participants used in the experiment was conventionally lateralized in brain organization. The results of these tests also demonstrate that the equipment was functioning appropriately and that a test of the cerebral lateralization ESP hypothesis was feasible.

EEG ANOVA analyses

The significant main effect difference of state (i.e., the production of significantly more left hemisphere and right hemisphere alpha in the ESP condition as opposed to the sensory condition) was likely the result of differences in the sensory and ESP tasks. In the ESP condition, participants were required to mentate internally, while focussing their eyes upon a lit but imageless TV screen. Aural input (TV white noise) was also monotonous, and the general lack of diverse sensory stimulation during the ESP condition may account entirely for the significant increase in overall alpha production during the ESP condition. This interpretation is consistent with some recent evidence provided by Ray and Cole (1985) who found that alpha abundance was greater during "inner imagining" tasks than during tasks which required the intaking of information from the environment.

The significant tape x hemisphere x sex interaction in males´ and females´ repeated measures ANOVA appears, at face value, to confirm the hypothesis of congruence in the lateralized processing of sensory and ESP information -with male dextrals showing the predicted cerebral lateralization effect to a significantly greater extent than female dextrals. However the confirmation of the hypothesis is not robustly supported in the finer analysis of relevant cross cuts of the data. Circumspection should, of course, attend the assessment of the internal tests because of their reduced sample sizes, but females, in fact, showed a tendency to be more ´right hemispheric´ for the City Tape than for the Love Tape - a "reversal" which had not been anticipated. With hindsight it is evident that, given the marginal lateralization effect of males´ hemispheres for the two stimulus videotapes, females´ hemispheres would have had to have exhibited the opposite lateralization effect for the test of the hypothesis (i.e., the tape x hemisphere x sex interaction) to have been significant. Females´ modest reversal of brain activation patterns for the stimulus videotapes (only suggestive at $p<.09$ for the total body of females´ data) probably requires no interpretation. Nevertheless it is tempting to conjecture that females, as a group, may be more analytical than males when evaluating or responding to emotional disclosures made by others.

EEG Correlational Analyses

The result of the correlation of differences in lateralized processing for two sets of stimuli in sensory and ESP conditions confirms the hypothesis of congruence in the asymmetrical processing of cognitive material during sensory and ESP perception. The data, which were significant for the group of subjects, imply that congruence is a meaningful concept when applied to hemispheric specialization effects during sensory and ESP processing. They support the conclusion that during extrasensory perception, brain activation is typical, rather than atypical, of brain activation during comparable sensory stimulation.

It is particularly encouraging that the cognitive and physiological data in this experiment provide two independently measured but converging lines of evidence for extrasensory processing. Not only were consciously reported memories of attempted ESP mentations matched to ESP stimulus videotapes more accurately than chance would predict (in the absence of any ESP processing), but also participants' differences in unconscious patterns of relative hemispheric engagement while actually viewing the two videotapes correlated significantly with their hemispheric differences in unconscious brain activity while attempting to respond to the videotapes extrasensorially.

When we select the participants whose mentations gave the best evidence for psi (i.e., participants whose mentations were correctly matched to the videotapes by themselves and two independent judges in a final judging) we find that their independently significant positive brain correlation is also significantly higher than that of the participants whose mentations were not correctly matched to the videotapes by all judges. Hemisphere data for the group of participants whose attempted ESP mentations showed a mixed pattern of judging, or whose mentations were incorrectly matched to the videotapes by all judges, rendered a negative correlation which was close to pure chance. Therefore we observe that the individuals who gave the best evidence for psi processing in the mental domain of consciously remembered perceptions are the same individuals who provided the best physiological evidence for psi processing in the physical domain of unconscious hemispheric activity. In short, these data demonstrate converging lines of evidence for psi from physical and mental domains of consciousness.

This conclusion is not without controversy. A question has been raised about the legitimacy of performing "... a separate correlation of the laterality scores of subjects who produced the strongest ESP scores ... [because] ... as the mentations of these subjects had already been judged to bear a similarity to the target, they might be expected to manifest similar brain patterns during the ESP condition and when they were actually viewing the target." The critic (note 4) suggests here that the converging lines of evidence provide no new information -that is, that they merely reflect two sides of the same coin. The objection presumes that the physiological measurements of subtle differences in unconscious target-related lateralized processing during two clairvoyance attempts were so refined as to mirror the cognitive differences in participants' subsequent conscious reports -reports which later prove to be discriminated on the basis of the same differences which were reflected in the brain measurements.

While this appears to be a feasible interpretation of the convergence, it is not well supported by the data. It is unlikely that overt similarity between the mentations and their respective target videotapes is a satisfactory explanation for the significant difference in correlations for the successfully and unsuccessfully judged participants. Short of this, and perhaps more parsimoniously, it can only be adjudged that the significantly higher correlation for successfully judged participants provides evidence that judges' successful matches were concordant with something that was also happening in the brains of the participants.

The test of correlated asymmetry differences between the successfully and unsuccessfully judged participants is crucial, moreover, because without the test we would have no way of knowing if the group of unsuccessfully judged participants had actually contributed more in the way of concordant processing to the success of the overall group correlation than had the group of successfully judged participants. Another possibility, which the test resolves, is that there may have been no difference in the amount of concordant processing that each group (the successfully and the unsuccessfully judged) contributed to the success of the overall correlation.

The test also allows us to discredit any conjecture that the judges' successful matches were due merely to judges' lucky guesses. For there are several ways that the mentation reports could have been successfully matched to the videotape targets on merits other than those implying 20 minutes of corresponding imagery.

One way is that judges' psi (or statistical anomaly for that matter) could have been operating to produce the correct matches. If this were true, we would not expect a significant correlation of corresponding differences in brain activity for the two videotapes to be exhibited between the sensory and ESP trials. Nor would we expect the correlation of differences in brain activity to be significantly higher for the successfully judged participants than for the unsuccessfully judged ones.

Another way that participants' mentations could have been successfully matched to the videotapes (whether or not ESP occurred) is that judges, when evaluating mentations, could have been swayed in favor of a correct matching by one or two isolated comments (perhaps even a single comment out of several pages of comments). Even if ESP were responsible for such fragmentary episodes of ideation (with their partial epochs of corresponding brain activity) we would not expect a relationship between videotape differences in brain activity for the tapes in the sensory and ESP trials because the successful matches would have been derived from only a very small portion of the overall brain activity.

A further basis for the successful matchings of the mentations to the target videotapes (an interpretation which is implied by the question of legitimacy in the difference test of correlations for successfully and unsuccessfully judged participants) could have been that the successfully judged mentations were so similar to their respective ESP targets that participants' brain activity differences when watching the two videotapes correlated with their brain activity differences when attempting to fathom the contents of the videotapes with psi. While this represents the ideal outcome of the study, it is surely an unwarranted conclusion, for in fact the psi mentations were in many respects overtly dissimilar to the videotapes. The lackluster results in the earliest stage of judging argue against any 'gross similarity' hypothesis. Recall that judges found participants' mentations to be as 'similar' to the wrong tapes as they did to the right tapes in the initial stage of judging. Moreover there is evidence from judges' introspective comments that judges often differed in what they judged to be salient, or similar, in the reports.

But if the successful blind judging results were not due to manifestly similar ESP ideations (which artifactually produced a significantly higher brain correlation for the successfully judged participants than for the unsuccessfully judged ones) what were they due to and why was the ESP pattern reflected in the unconscious (EEG) data? David Bohm (1986), in his Gardner Murphy Memorial lecture 'A new theory of the relationship of mind and matter' provides a notion of "meaning" as the link between that which is mental and physical in nature. "This link is indivisible, in the sense that information contained in thought, which we feel to be on the 'mental' side, is at the same time a neurophysiological, chemical, and physical activity, which is clearly what is meant by this thought on the 'material' side" (Bohm, 1986, p. 128). According to his theory, extremely subtle levels of mental activity -which incorporate this notion of meaning- can organize the range of information content embodied in our thoughts into a single greater whole. "There is never any real division between mental and material sides at any stage of the overall process" (Bohm, 1986, p. 129). This is because, at the quantum level, the individual "electron dance" is "orchestrated" by the meaning embodied in increasingly higher order or super wavefunctions.

In a parapsychological context, contact between an observer or participant and the target (in this case a remote videotape) would "... depend more on similarity or 'resonance' of meanings than on location in space." (Bohm, 1986, p. 132). The fertile intelligence of the observer, once meaning is apprehended (on whatever level of sublety), gives shape to the play of thoughts which implicitly manifests the similarity or resonance of meaning. Simultaneously, on quantum levels, "... the 'dance' of the electrons ... the objective meaning of the information content in the 'score' of the wavefunction" executes the physiologically interwoven meaning (Bohm, 1986, p. 126).

Bohm's model appears to afford a comfortable and natural explanation for the covert similarity observed between participants' conscious and unconscious responses and the remote videotape target. Extrapolating from the model, it can be conjectured that the observer (i.e., the participant) and the observed (the target videotape), have been united by the meaning implicit in the resonance between them. Each observer (to the extent that this meaning has been penetrated) will manifest the similarity both physically, in the play of quantum energy (resulting in correlated macroscopic EEG measurements) and mentally in the shape of the individual´s unique cognitive associations (which can be successfully matched to the videotape target.) Therefore, although the thoughts appear disparate from the actual target object, both mental and physical indices can reflect similarity to the meaning implicit in the target (note 5). The cognitive and physical indices expound a common meaning because both indices are reflections of what, in essence, is one and the same consciousness.

While it is premature to speculate on how well Bohm´s model satisfies the range of observed phenomena in psi research, it can clearly accomodate (in theory at least) a wide spectrum of individual differences in free-response psi responding. In the context of the present experiment, the model addresses the crucial problem of overt dissimilarity between the targets and responses. The empirical outcome of the experiment is thereby given a theoretical coherence which is both elegant and intuitively plausible.

ABSTRACT

EEG alpha was monitored from bilateral temporal-parietal sites while 10 male and 10 female dextrals attempted to use ESP to fathom the contents of two videotapes which were played, in random order, in a distant room. After each ESP session, participants reported their mentations. Subsequently, EEG alpha was monitored while participants viewed the videotapes (again randomly sequenced) on a TV monitor. One of the videotapes contained material with emotion inducing properties, while the other videotape was generally devoid of emotion inducing content. A repeated measures ANOVA on the sensory and ESP data gave a significant tape x hemisphere x sex interaction ($p<.05$), with males´ data demonstrating a greater relative right hemispheric participation in processing the emotional as opposed to the nonemotional material (as had been predicted) and females´ data showing the converse hemispheric pattern (a ´reversal´ which had not been fully anticipated). When participants attempted to match their mentations to the videotapes, results were at chance, but one of two outside blind judges exhibited extrachance accuracy in a final matching of the mentation reports to the videotapes ($p<.007$, one-tailed). Differences in individuals´ mean laterality indices for processing the two videotapes in the sensory condition were significantly correlated with their respective differences in processing the two videotapes in the

ESP condition ($p<.04$, one-tailed). A separate correlation for nine participants whose mentations were correctly matched to the videotapes by all judges (the participants themselves and both outside judges) was both independently significant ($p<.015$, one-tailed) and significantly higher ($p<.05$, one-tailed) than the asymmetry correlation for the participants whose mentations had shown a mixed pattern of judging, or whose mentations had been incorrectly matched to the videotapes by all judges. The data demonstrate converging lines of evidence for conscious and unconscious ESP.

NOTES

1) When corrected for selection, participants' handedness; eye-dominance; thumb-dominance; and EPQ variables were not significantly correlated with brain variables or ESP scores. Space considerations prevent the full reporting of these analyses in this paper.
2) Prior to the computation of the Pearson correlation coefficient, a Spearman correlational analysis had been performed on the same data. The resulting coefficient was $rs=.5579$; $p<.007$, one-tailed, for the sample of 19 participants. However, because the scores represent interval level data, a Pearson correlational analysis is reported.
3) A Spearman analysis for the group of participants whose mentations were correctly matched by all judges gives $rs=.8667$; $p<.005$, one-tailed. The correlation for those whose mentations showed a mixed pattern of judging, or whose mentations were incorrectly matched by all judges gives $rs=.3455$; ns.
4) A prior referee who has declined to be identified.
5) According to Bohm (1986, p. 132) this interpretation presupposes that the quantum mechanical wavefunction satisfying Schrodinger's equation has been modified, creating a drift of particles in the direction of the implicit 'meaning'.

ACKNOWLEDGEMENTS

This paper is based on dissertation research submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the City University of New York.

Grateful appreciation is due to my experimental assistant, Kevin

Bjorn Reese; to my dissertation advisors Dr. Gertrude Schmeidler, Dr. Howard Ehrlichman, and Dr. Louis Gerstman; to Wendy Clark and Hilary Harris for the loan of ´The Love Tapes´ and ´Organism´; to Dr. Majorie Wiener and the late Jack Barrett for expert advice and assistance with alpha monitoring; and to the participants and judges who demonstrated, once again, that psi is possible.

Requests for reprints may be sent to the author´s present affiliation: Psychology Department, Room 303; The New School for Social Research; 66 West 12th Street; New York, NY 10011. U.S.A..

REFERENCES

Broughton, R.S. ´Brain hemisphere specialization and its possible effects on ESP performance´, in J.D. Morris, W.G. Roll and R.L. Morris (eds.): Research in Parapsychology 1975, Metuchen, NJ, Scarecrow Press, 1976, 98-102.

Broughton, R.S. ´Brain hemisphere differences in psi influenced reaction time´, in J.D. Morris, W.G. Roll, and R.L. Morris (eds.): Research in Parapsychology 1976, Metuchen, NJ, Scarecrow Press, 1977, 86-90.

Bohm, D.J. ´A new theory of the relationship of mind and matter, J. A.S.P.R., 80, 1986, 113-135.

Carmon, A. and Nachshon, I. ´Ear asymmetry in perception of emotional non-verbal stimuli´, Acta Psychologica, 37, 1973, 351-357.

Charman, D.K. ´Do different personalities have different hemispheric asymmetries?´, A brief communique of an initial experiment, Cortex, 15, 1979, 655-658.

Davidson, E.D. and Ehrlichman, H. ´Lateralized cognitive processes and the electroencephalogram´, Science, 207, 1980, 1005-1006.

Dean, E.D. ´Long-distance plethysmograph telepathy with agent under water´, Proceedings of the Parapsychological Association, 6, 1969, 41-42.

Dean, E.D. and Nash, C.B. ´Coincident plethysmograph results under controlled conditions´, J. A.S.P.R., 44, 1967, 1-13.

Dimond, S.J. and Farington, L. ´Emotional response to films shown to the right or left hemisphere of the brain measured by heart rate´, Acta Psychologica, 41, 1977, 255-260.

Ehrlichman, H. and Wiener, M.S. ´Consistency of task-related EEG asymmetries´, Psychophysiology, 16, 1980, 247-252.

Galin, D. and Ornstein, R. ´Lateral specialization of cognitive mode: An EEG study´, Psychophysiology, 9, 1972, 412-413.

Gruber, F. and Segalowitz S. ´Some issues and methods in the neuropsychology of language´, in S. Segalowitz and F. Gruber (eds.): Language Development and Neurological Theory, New York, Academic Press, 1977.

Haggard, M.P. and Parkinson, A.M. ´Stimulus and task factors as determinants of ear advantages´, Quarterly Journal of Experimental Psychology, 23, 1971, 168-177.

Hansch, E.C. and Pirozzollo, F.J. ´Task relevant effects on the assessment of cerebral specialization for facial emotion´, Brain and Language, 10, 1980, 51-59.

Haraldsson, E. ´Psychological variables in a GESP test using plethysmograph recordings´, Proceedings of the Parapsychological Association, 7, 1970, 6-7.

Hardyck, C. and Petrinovich, L.F. ´Left handedness´, Psychological Bulletin, 84, 1977, 385-404.

Hartwell, J.W. ´Contigent negative variation as an index of precognitive information´, E.J.P., 2, 1978, 83-103.

Hearne, K.M.T. ´Visually evoked responses and ESP´, J. S.P.R., 49, 1977, 648-657.

Hearne, K.M.T. ´The effect on the subject (in SWS and REM state) of electric shocks to the agent: an ESP experiment´, J. S.P.R., 51, 1981a, 87-92.

Hearne, K.M.T. ´Visually evoked responses and "ESP": failure to replicate previous findings´, J. S.P.R., 51, 1981b, 145-147.

Heilman, K.M., Scholes, R. and Watson, R.T. ´Auditory affective agnosia: Disturbed comprehension of affective speech´, Journal of Neurology, Neurosurgery, Psychiatry, 38, 1975, 69-72.

Hughlings-Jackson, J. ´On affections of speech from diseases of the brain´, Brain, 2, 1879, 203-222.

Kelly, E.F. and Lenz, J.E. ´EEG changes correlated with a remote stroboscopic stimulus: a preliminary study´, in J.D. Morris, W.G. Roll and R.L. Morris (eds): Research in Parapsychology 1975, Metuchen, NJ, Scarecrow Press, 1976, 58-63.

Kennedy, J.E. ´Possible paranormal components of anticipation: Psychophysiological explorations´, Abstract, J.o.P., 43, 1979, 360-361.

King, F.L. and Kimura, D. ´Left-ear superiority in dichotic perception of vocal nonverbal sounds´, Canadian Journal of Psychology, 20, 1972, 111-116.

Kinsbourne, M. ´The cerebral basis of lateral asymmetries in attention, Acta Psychologica, 33, 1970, 193-201.

Levin, J. and Kennedy, J. ´The relationship of slow cortical potentials to psi information in man´ (abstract), J.o.P., 39, 1975, 25-26.

Ley, R.G. and Brayden, M.P. ´Hemispheric differences in processing emotions and faces´, Brain and Language, 7, 1979, 127-138.

Lloyd, D.H. (pseudonym) ´Objective events in the brain correlating with psychic phenomena´, New Horizons, 1, 1973, 69-75.

Luria, A.R. ´The working brain´, New York, Basic Books, 1973.

Maher, M., Peratsakis, D. and Schmeidler, G.R. ´Cerebral lateralization effects in ESP processing: An attempted replication´, J. A.S.P.R., 73, 1979, 167-177.

Maher, M. and Schmeidler, G.R. ´Quantitative investigation of a recurrent apparition´, J. A.S.P.R., 69, 1975, 341-352.

Maher, M., and Schmeidler, G.R. ´Cerebral lateralization effects in ESP processing´, J. A.S.P.R., 71, 1977, 261-271.

Masters, R. and Houston, J. ´Mind games´, New York, The Viking Press, 1972, 8-9.

May, E.C., Targ, R. and Puthoff, H.E. ´EEG correlates to remote light flashes under conditions of sensory shielding´, in C.T. Tart, H.E. Puthoff, and R. Targ (eds.): Mind at Large, New York, Preager, 1979, 127-136.

McGlone, J. ´Sex differences in human brain asymmetry: A critical survey´, The Behavioral and Brain Sciences, 3, 1980, 215-263.

McKee, G., Humphrey, B. and McAdam, D.W. ´´Scaled lateralization of alpha activity during linguistic and musical tasks´, Psychophysiology, 10, 1973, 441-443.

Millar, B. ´An attempted validation of the "Lloyd effect", in J.D. Morris, W.G. Roll, and R.L. Morris (eds.):Research in Parapsychology 1975, Metuchen, NJ, Scarecrow Press, 1976, 25-27.

Millar, B. ´Physiological detectors of psi´, E.J.P., 2, 1979, 456-478.

Moore, W.H. and Haynes, W.O. ´A study of alpha hemispheric asymmetries for verbal and nonverbal stimuli in males and females´, Brain and Language, 9, 1980, 338-349.

Moss, T and Schmeidler, G.R. ´Quantitative investigation of a "haunted house" with sensitives and a control group´, J. A.S.P.R., 62, 1968, 399-410.

Oppenheimer, J.M. ´Studies of brain asymmetry: Historical perspective´, in S.J. Dimond and D.A. Blizard (eds.): Evolution and Lateralization of the brain, New York, New York Academy of Sciences, 1977, 4-17.

Ray, W.J. and Cole, H.W. ´EEG Alpha activity reflects attentional demands and Beta activity reflects emotional and cognitive processes´, Science, 228, 1985, 750-752.

Rhine, L. ´Research methods with spontaneous cases´, in B.B. Wolman (ed.): Handbook of Parapsychology, New York, Van Nostrand Reinhold, 1977, 59-80.

Ross, E.D. ´The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere´, Archives of Neurology, 38, 1981, 144-148.

Ross, E.D. and Mesulam, M. ´Dominant language functions of the right hemisphere´, Archives of Neurology, 36, 1979, 144-145.

Safer, M.A. and Leventhal, H. ´Ear differences in evaluating emotional tones of voice and verbal content´, Journal of Experimental Psychology: Human Perception and Performance, 3, 1977, 75-82.

Schmeidler, G.R. ´Quantitative investigation of a "haunted house"´, J. A.S.P.R., 60, 1966, 139-149.

Schouten, S.A. ´Autonomic psychophysiological reactions to sensory and emotive stimuli in a psi experiment´, E.J.P., 2, 1976, 456-478.

Sidtis, J.J. ´On the nature of the cortical function underlying right hemisphere auditory perception´, Neuropsychologia, 18, 1980, 321-330.

Suberi, M. and McKeever, W.F. ´Differential right hemispheric memory of emotional and nonemotional faces´, Neuropsychologia, 15, 1977, 757-768.

Targ, R. and Puthoff, H. ´Information transmission under conditions of sensory shielding´, Nature, 251, 1974, 602-607.

Targ, R. and Puthoff, H. ´Mind reach´, New York, Delacorte Press, 1977.

Tart, C.T. ´Physiological correlates of psi cognition´, International Journal of Parapsychology, 5, 1963, 375-386.

Tart, C.T. ´Physiological correlates of psi cognition: Some methodological considerations´, in W.G. Roll, J. Beloff, and R.A. White (eds.): Research in Parapsychology 1982, Metuchen, NJ, Scarecrow Press, 1983, 221-223.

Trotman, S. and Hammond, G. ´Sex differences in task dependent EEG asymmetries´, Psychophysiology, 16, 1979, 429-431.

Willis, S.G., Wheatley, G.H. and Mitchell, O.R. ´Cerebral processing of spatial and verbal-analytic tasks: An EEG study´, Neuropsychologia, 17, 1979, 473-484.

Witelson, S.F. ´Sex and the single hemisphere: Specialization of the right hemisphere for spatial processing´, Science, 193, 1976, 425-427.

Michaeleen Maher
Psychology Department, Room 303
The New School for Social Research
66 West 12th Street
New York, NY 10011.
U.S.A.

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART V. THE GROUP OF CONTROL SERIES WITH NON-PSYCHICS

Hendrik G. Boerenkamp
University of Utrecht

In a previous paper the design of an experimental study of paranormal impressions of psychics was described (Boerenkamp, 1985a). Psychics are defined as persons who believe themselves able to obtain paranormal impressions at will. Usually, psychics are consulted by clients about problems related to themselves and to persons in their environment. The entire study consists of a predetermined number of series of sessions with a group of 12 psychics and two groups of 12 non-psychics in which the participants gave their impressions about persons unknown to them.
A series of sessions means that each psychic or non-psychic gave his or her impressions about one target person. The series with the psychics were divided in three standard series, plus two subgroups of five experimental series each. Two of the three standard series were held at the start of the investigation (standard series A and B). These standard series were included to provide a description of the content of sessions when psychics apply their assumed psi abilities under conditions which resemble as much as possible the daily circumstances of sessions with clients. For the results, see Boerenkamp, 1985b. A further aim of including these standard series was to compare the content of these sessions with the content of the sessions of each of the experimental series, which were held to study the effect of a number of variables on the content of the sessions, i.e. the statements of the psychics about target persons. The results of the experimental series were discussed in Boerenkamp, 1985c and

1986a.

In order to compare statements provided by psychics with statements provided by non-psychics under similar conditions two control groups were created. The first one consisted of a group of 12 persons matching the group of psychics in age, sex and level of education ('matched' control group). This group was formed by randomly approaching persons of the desired age, sex and level of education in different districts of the town of Utrecht with the request to participate in an investigation of 'knowledge of people'. The second group was not matched for age, sex and level of education but composed of 12 persons (doctors, psychologists and lawyers) who, because of their profession, were experienced in dealing with problems of people ('professional' control group). They were also approached with the request to participate in the same type of investigation. Both control groups carried out one standard series (in which a photograph and an object of the target person were presented) and one experimental series (i.e. series 1 in which only an object of the target person was presented). Hence, four control series were run.

In this paper the results of the four control series are discussed and compared with the results of the standard series and experimental series 1 with the psychics. (In this paper experimental series 1 will be referred to as the experimental series without further more).
No specific hypotheses were stated in advance. The four control series of sessions with the non-psychics were analyzed in the same way as the series of sessions with the psychics.

In the standard series the psychics and the non-psychics were presented with a photograph and an object belonging to a target person. In the experimental series the psychics and the non-psychics were presented with only an object of the target person.
In the series with the psychics two researchers (the author and S.S.) acted as the sitter, each of them in half of the sessions of each series. In the series with the two groups of non-psychics, two assistant researchers acted as the sitters. One of them (L.H.) acted as the sitter in the series with the 'matched' control group. The other (R.F.) acted as the sitter in the series with the 'professional' control group. Both assistant researchers were unacquainted with the results of the sessions with the psychics.
In the series with the psychics six different target persons of different age and sex were used in order to neutralize a possible systematic influence of age and sex of the target person (and

therefore in the description of the behaviour of the psychics). Target persons were chosen from the environment of the sitters (the author and S.S.). The six target persons were men and women of about 25, 45, and 65 years of age.
Two of these six target persons were used in the control series with the non-psychics. A woman of about 45 years of age and a man of about 65 years of age. The two target persons were allocated in a systematic way to the 12 non-psychics in the ´matched´ as well as in the ´professional´ control group. The first half of each control group made statements about the woman and the second half about the man in the standard series. In the experimental series the first half of each control group made statements about the man and the second half about the woman. Each non-psychic did first the session of the standard series and then the session of the experimental series.

In all series with the psychics they were invited to give paranormal impressions concerning the target person. In all series with the non-psychics they were invited to give ´uninhibited´ impressions concerning the target person. This was accomplished by suggesting them that they possibly ´knew´ more about the person than they were aware of. They received no further instructions. Nothing was told about the investigation of paranormal impressions of psychics.
In all series both psychics and non-psychics were told that the sitter was acquainted with the target person, but they were not informed beforehand about special problems in the life of the target person. Actually, in the control series the sitters were not acquainted with the two target persons. Instead, they had a list with information concerning the target person. The list contained some information concerning physical characteristics, psychological characteristics, relations with other people, civil status, circumstances in work, etcetera.

In all series the psychics and the non-psychics received immediate feedback to their statements in the form of an affirmation or a denial in the form of "yes", "no" or "I don´t know" (confined informative action), occasionally followed by some clarification which provided additional related information (extended informative action). No information unrelated to the topic being discussed was provided in the feedback.
The sitters in the sessions of the control series were instructed to balance as much as possible the number of ´confined´ and ´extended´ informative actions, in accordance with the behaviour of the sitters in the sessions with the psychics.

The sessions are described in terms of the number of statements of the psychics and non-psychics and the number of informative actions (feedback) of the sitters. Also the results of the informational, structural and interactional analyses are presented. In the informational analysis it is established which of the psychic's and non-psychic's statements are sufficiently specific and spontaneous to assign them potential paranormal value, taking into account the available information at the moment the statement was made. Statements having potential paranormal value which turn out to be correct are called statements with positive paranormal value. In the structural analysis various characteristics of the set of statements are studied, for example the topics discussed in the statements. The interactional analysis concerns the actions psychics and non-psychics take when the sitter denies the correctness of the content of a statement, for example giving another interpretation of the impression. For a more detailed discussion of these concepts, see Boerenkamp, 1985a.

The two standard series A and B were carried out with 12 psychics. One standard series was carried out with 12 non-psychics in each control group. From the previous study (Boerenkamp, 1985b) it appeared that the behaviour of the psychics in standard series A was very consistent with their behaviour in standard series B. Hence, the data of standard series A and B were combined in order to minimize the role of accidental influences on the content of the standard series with psychics. The combined data of the two standard series were divided by two and then compared with the data of the standard series with each of the control groups.
The experimental series (series 1) was conducted with 11 of the 12 psychics (the contribution of one psychic was lacking because he died) and therefore the data of the experimental series with the 11 psychics are compared with the data of each of the experimental series with 11 non-psychics in each control group.

As stated above, the non-psychics were approached with the request to participate in an investigation of 'knowledge of people', i.e. how well they are able to judge people unknown to them, and to give their 'uninhibited' impressions of the person who is pictured in the photograph and/or to whom the object belongs. In fact, apart from the labeling this amounts to the same as psychics do. However, because of the rather uncommon nature of the request, it can be expected that compared to psychics non-psychics make fewer statements. After all, common people are not used to give impressions about people unknown to them by using only a photograph or an object. In order to balance against this tendency in these series the sitter asked specifically for impressions about relations with others, civil status and circumstances in work in the case the subject did not volunteer such impressions.

Because the judges of the sessions with the non-psychics could notice from the transcripts that non-psychics were acting as psychics the judges were told that the question was which statements of non-psychics, acting as psychics, might have potential paranormal value. They were told that non-psychics might have paranormal impressions without being aware of it. The judges of the statements of the psychics were not aware of the results of the procedure on the statements of the non-psychics and vice versa.

The six series involved in this paper are denoted by:
PSY-S: standard series with psychics (N=12)
MNP-S: standard series with ´matched´ non-psychics (N=12)
PNP-S: standard series with ´professional´ non-psychics (N=12)
PSY-E: experimental series with psychics (N=11)
MNP-E: experimental series with ´matched´ non-psychics (N=11)
PNP-E: experimental series with ´professional´ non-psychics (N=11)

Number of statements and informative actions

The number of statements made by the psychics and the non-psychics and the number of informative actions (feedback) by the sitters in the different series is presented in Table 1.

The psychics made much more statements than the non-psychics in the standard series as well as in the experimental series (Mann-Whitney U Test: U varies from U=1.00, $p<.002$ for the PSY-S vs the PNP-S series to U=6.00, $p<.002$ for the PSY-E vs the MNP-E series; all Mann-Whitney U tests applied in this paper are two-tailed; table K of Siegel´s Nonparametric Statistics, McGraw-Hill, 1956, is consulted to find the significance level of the observed U-value; the Mann-Whitney U Test is further denoted as MWU). However, the behaviour of the sitters might also have contributed to this difference. The sitters in the sessions with the psychics gave relatively more feedback responses to the psychics than the sitters in the sessions with the non-psychics in the standard series as well as in the experimental series (MWU: U varies from U=13.50, $p<.002$ for the PSY-S vs the MNP-S series to U=19.00,

TABLE 1
Number of statements by the psychics and non-psychics and number of informative actions by the sitter

	PSY-S N=12	MNP-S N=12	PNP-S N=12	PSY-E N=11	MNP-E N=11	PNP-E N=11
statements (N)	1060	310	254	798	288	190
range	39-171	16-46	11-38	29-165	14-48	7-27
mean	88	26	21	73	26	17
informative actions (N)	720	126	123	562	144	77
con. info. actions (N)	436	74	83	352	94	51
ext. info. actions (N)	284	52	40	210	50	26
info.act./statements	68%	41%	48%	70%	50%	41%
ext.info.act./info.act.	39%	41%	33%	37%	35%	34%

p<.02 for the PSY-E vs the MNP-E series). This might be due to the fact that the sitter in the sessions with the non-psychics were not acquainted with the target person. Furthermore, the sitters asked for impressions about certain topics (relations with others, civil status and circumstances in work) in the case the non-psychic did not volunteer such impressions. This active behaviour (asking questions) might have suppressed to some extent the reactive behaviour (giving feedback) of the sitter. On the other hand, the number of statements by the non-psychics certainly would have been still lower if the sitter had not asked for such impressions.
In all series with psychics and non-psychics the sitters gave extended feedback responses in about the same degree compared to confined feedback responses (MWU: U varies from U=50.00, n.s. for the PSY-S vs the PNP-S series to U=55.50, n.s. for the PSY-E vs the MNP-E series).

In the standard series the non-psychics in the ´matched´ control group (matched in age, sex and level of education) did not make significantly more statements than the non-psychics in the ´professional´ control group (doctors, psychologists, laywers) (MWU: U=49.00, n.s.). However, in the experimental series the ´matched´ non-psychics made more statements than the ´professional´ non-psychics (MWU: U=29.00, p<.05). This difference can not be attributed to differences in feedback by the sitter (MWU: U=49.50, n.s.). The 'professional' non-psychics appeared more 'inhibited' in giving impressions than the 'matched' non-psychics in the case that only an object belonging to the target person was presented.

As expected, in the sessions with psychics the number of informative actions made by the sitter depended on the number of statements made by the psychic (standard series: rs=.71, t=4.83, df=22, p<.001; experimental series: rs=.86, t=5.01, df=9, p<.001). The same but less significant pattern applied in the sessions with the non-psychics (MNP-S series: rs=.67, t=2.82, df=10, p<.02; PNP-S series: rs=.65, t=2.69, df=10, p<.05; MNP-E series: rs=.62, t=2.34, df=9, p<.05; PNP-E series: rs=.73, t=3.24, df=9, p<.02).
In the case of the sessions with the psychics it seems justified to conclude that the number of informative actions by the sitter depended on the number of statements made by the sitter as the psychic had the initiative in the session. In the case of the sessions with the non-psychics it may not be justified to interpret the (less significant) correlations in this way. In these sessions it was much less clear who had the initiative, as the non-psychics were not used to the task. In sessions with non-psychics we only can say that the number of statements and the number of informative actions are related.

The informational analysis

The first part of the informational analysis of each session, which consisted of the itemization of the statements and of the estimation of the potential paranormal value of each statement, was carried out by two judges. For each series of sessions a different group of judges was employed. This part of the procedure is described in detail in Boerenkamp, 1984. The selection was based on the estimate of the probability of correspondence (specific versus vague) combined with the estimate of the degree of spontaneity (spontaneous versus inferred) of each statement by two judges. Thus, each statement was rated by two judges in two different ways on a four-point scale. On these scales a low degree of potential paranormal value is represented by a score of 1 or 2 and a high degree of potential paranormal value by a score of 3 or 4. In the study mentioned above it was found that two judges agreed on 82% of the statements to which of the categories (1 or 2 versus 3 or 4) it should be assigned.

The inter-rater-reliability for two judges in each of the four control series was comparable to the inter-rater-reliability obtained in the previous study. In the MNP-S series the judges agreed on 85% of the statements, in the PNP-S series on 85%, in the MNP-E series on 83% and in the PNP-E series on 85% of the statements. In the PSY-S series this figure was 84% and in the PSY-E series 81%. The distributions of the scores of potential paranormal value, based on the combined scores of two judges rating each statement on both probability of correspondence and degree of spontaneity on scale ranges from 1 to 4, are presented in Table 2.

TABLE 2
Distribution of scores of potential paranormal value in the series with psychics and non-psychics

	Categories of potential paranormal value									
	low	(%)	low-med	(%)	medium	(%)	med-high	(%)	high	(%)
PSY-S series	587	(55)	239	(23)	140	(13)	61	(6)	33	(3)
MNP-S series	183	(59)	69	(22)	39	(13)	16	(5)	3	(1)
PNP-S series	172	(68)	33	(13)	28	(11)	12	(5)	9	(4)
PSY-E series	426	(53)	164	(21)	129	(16)	49	(6)	30	(4)
MNP-E series	133	(46)	79	(27)	48	(17)	18	(6)	10	(3)
PNP-E series	110	(58)	38	(20)	22	(12)	12	(6)	8	(4)

Applying a cut-off criterion between the medium and medium-high categories, each of the six series yielded about 9% (range 6% to 10%) statements with potential paranormal value (ppv).

According to the Kolmogorov-Smirnov two-sample test the distributions of scores of ppv in the PSY-S and MNP-S series do not significantly differ (Dmax=.04, chi-square=1.28, df=2, n.s.; all Kolmogorov-Smirnov two-sample tests in this paper are two-tailed; the Kolmogorov-Smirnov two-sample test is further denoted as KS). The ´matched´ non-psychics compared to the psychics made relatively the same percentage of statements with ppv (MWU: U=60.50, n.s.). The PSY-S and PNP-S series significantly differ (Dmax=.12, chi-square=12.48, df=2, p<.01). However, the difference between the two series is attributable to differences between the lowest categories and not to

differences between the relevant medium-high and high categories. The 'professional' non-psychics compared to the psychics made relatively the same percentage of statements with ppv (MWU: U=57.00, n.s.). The distributions of scores of ppv in the MNP-S and the PNP-S series do not significantly differ (KS: Dmax=.09, chi-square=4.21, df=2, n.s.). The 'matched' non-psychics compared to the 'professional' non-psychics made relatively the same percentage of statements with ppv (MWU: U=69.00, n.s.).
The distribution of scores in the PSY-E series is not significantly different from the distribution of scores in the MNP-E series (KS: Dmax=.07, chi-square=4.39, df=2, n.s.) and in the PNP-E series (KS: Dmax=.05, chi-square=1.25, df=2, n.s). The non-psychics compared to the psychics made relatively the same number of statements with ppv (MWU: PSY-E vs MNP-E: U= 54.00, n.s.; PSY-E vs PNP-E: U=54.00, n.s.). However, the distributions of scores in the MNP-E and PNP-E series differ to a significant degree (KS: Dmax=.12, chi-square=6.28, df=2, p<.05), but also in this case the difference between the two series is attributable to differences between the lowest categories and not to differences between the relevant medium-high and high categories. The 'professional' non-psychics compared to the 'matched' non-psychics made relatively the same percentage of statements with ppv (MWU: U=50.00, n.s.).

Thus, the psychics compared to the non-psychics did not take more 'risk' by making more spontaneous and specific statements. Or one might say that non-psychics, when asked for 'uninhibited' impressions about people unknown to them, dare to take as much 'risk' as psychics do.

From the data presented one might gain the impression that the 'matched' non-psychics responded in the same way as the 'professional' non-psychics did. In fact, in some respects this seems not to be the case. Although both groups of non-psychics in contrast to the psychics made quite a lot of comments on their more spontaneous and specific statements such as 'How dare I to state this?' especially in the experimental series, the 'matched' non-psychics seemed to take their impressions more seriously (believed in the potential correspondence between the content of their statements and the facts concerning the target person) than the 'professional' non-psychics. The behaviour of the 'matched' non-psychics seemed in this respect more similar to the behaviour of the psychics than the behaviour of the 'professional' non-psychics. However, it has to be noted that this is a subjective impression of the author.

The higher number of statements provided by the ´matched´ non-psychics in the case that only an object belonging to the target person was presented, can be considered one indication. A second indication might be that all ´matched´ non-psychics seemed to take the task seriously, while two ´professional´ non-psychics did not take the task seriously. They expressed that by making an explicit reference to the work of psychics ("I am not a clairvoyant but I will try to be one") followed by some very specific statements.

Of the 94 statements with ppv of the PSY-S series, 52 were made in the first half and 42 were made in the second half of the session. This difference is not significant. Comparable results were observed in the PSY-E series (46 in the first half and 33 in the second half). Thus, in each series more statements with ppv were made in the first half of the session. But, the difference within each series is not significant. The control series yielded a similar pattern (MNP-S series: 14 vs 5; PNP-S series: 14 vs 7; MNP-E series: 20 vs 8; PNP-E series: 10 vs 10).

The second step in the informational analysis was to establish how many of the 6 to 10% statements with ppv also fulfilled the criterion of ´sufficient degree of correspondence´. The statements with ppv were rated on a four point scale as being true or untrue by the researcher, who was acquainted with the target person. Statements which received a score of either 1 or 2 were called statements of negative paranormal value (statements rated as untrue: score 1, statements rated as probably untrue or more untrue than true: score 2) and statements which received a score of 3 or 4 were called statements of positive paranormal value (statements rated as more true than untrue or probably true: score 3, statements rated as true: score 4).
From Table 3 it appears that in all six series the degree of correspondence between the content of the statements with potential paranormal value and the facts about the person is low.

Only about 1% of the statements in each of the six series is spontaneous, specific and correct, i.e. has positive paranormal value. No significant differences are observed between the distributions of positive and negative paranormal statements in the different series. (KS: Dmax value varies from Dmax=0.10, chi-square=0.63, df=2, n.s. for the PSY-S vs the MNP-S series to Dmax=0.04, chi-square=0.28, df=2, n.s. for the PSY-S vs the PNP-S series). The psychics compared to the non-psychics were unable to make relatively more statements with positive paranormal value.

TABLE 3
Distribution of scores of 'degree of correspondence' of the statements with potential paranormal value in the series with psychics and non-psychics

	Categories of paranormal value							
	negative value (-pv)				positive value (+pv)			
score	1	(%)	2	(%)	3	(%)	4	(%)
PSY-S series	50	(53)	30	(32)	12	(13)	2	(2)
MNP-S series	12	(63)	4	(21)	2	(11)	1	(5)
PNP-S series	12	(57)	6	(29)	3	(14)	0	(0)
PSY-E series	39	(49)	25	(32)	12	(15)	3	(4)
MNP-E series	16	(57)	8	(29)	3	(11)	1	(4)
PNP-E series	10	(50)	6	(30)	2	(10)	2	(10)

The structural analysis

In each of the control series the number of statements with positive paranormal value was too low to render a meaningful statistical comparison between distributions of statements with positive paranormal value and statements with negative paranormal value as regards the different characteristics studied. But there appeared no indication that the statements with positive paranormal value are different with respect to any of the characteristics. Therefore, in the structural analysis of the control series only the set of statements with ppv and the set of statements without ppv were compared with respect to each characteristic.

(1) Topics discussed in the statements

The distribution of the topics in the six series is presented in Table 4. The distributions of all statements (all) and statements with ppv are presented in real numbers for the main categories (A-E), whereas for the subcategories (10-54) percentages are given. (For some categories the description is abbreviated because of the limited space in the table).

TABLE 4
Distribution of topics discussed in the statements

	PSY-S	MNP-S	PNP-S	PSY-E	MNP-E	PNP-E
Total number of topics	1428	413	335	1138	389	255
A Physical characteristics: all	231	79	66	344	117	64
Physical characteristics: ppv	36	1	4	40	14	6
10 Sex	0%	0%	0%	1%	3%	2%
11 Age	1%	4%	4%	3%	5%	9%
12 Appearance	5%	12%	14%	9%	18%	10%
13 Bodily health	8%	1%	2%	12%	2%	3%
14 Being alive or dead	2%	2%	0%	5%	2%	1%
B Psychological charact.: all	612	138	119	378	84	77
Psychological charact.: ppv	23	1	4	25	4	4
21 Personality traits	27%	30%	29%	19%	19%	21%
22 Psychological circumstances	14%	1%	3%	13%	1%	2%
23 Religious orientation	2%	2%	4%	1%	1%	7%
C Relations: all	260	49	48	143	47	17
Relations: ppv	13	2	2	12	0	1
31 Relations with family members	8%	4%	5%	7%	3%	2%
32 Relations with friends	7%	7%	9%	3%	8%	4%
33 Relation with sitter	3%	1%	0%	3%	1%	0%
D Specific topics: all	325	147	102	273	141	97
Specific topics: ppv	46	17	15	38	18	14
41 Civil status	3%	8%	4%	3%	8%	3%
42 Circumstances in work	9%	21%	17%	8%	15%	13%
43 Circumstances in living	3%	2%	1%	4%	3%	4%
44 Leisure activity	4%	2%	2%	3%	4%	9%
45 Specific name, event	4%	3%	6%	6%	7%	10%

The distributions of the topics in the series with the psychics were rather different from the distributions in the series with the non-psychics. The correlations between the frequencies of the topics (subcategories 10-45) are (with one exception) not significant (PNP-S and MNP-S series: rs=.39, t=1.58, df=14, n.s; PSY-S and PNP-S series: rs=.60, t=2.80, df=14, p<.02; PSY-E and MNP-E series: rs=.25, t=0.97, df=14, n.s.; PSY-E and PNP-E series: rs=.26, t=1.02, df=14, n.s.). The distributions of the topics in the series with each group of non-psychics were rather similar (MNP-S and PNP-S series: rs=.87, t=6.62, df=14, p<.001; MNP-E and PNP-E series: rs=.75, t=4.21, df=14, p<.001). However, it has to be noted that both groups of non-psychics were specifically asked for impressions about relations with others, civil status and circumstances in work. That must have contributed to the significance of the correlations between the distributions of the non-psychics and to the non-significance of the correlations between the distributions of the psychics and the non-psychics.

However, some spontaneous preferences of the non-psychics must be noted. In the standard series with the non-psychics the photograph induced quite a lot of statements about the age and appearance of the target person ("she is decently dressed", "he is about 50"). This means that in contrast to the psychics the non-psychics openly used the information in the photograph, probably in order to say at least something. On the other hand, in the experimental series in which only an object of the target person was presented, the non-psychics seemed to ´imagine´ age and appearance of the target person even relatively more often than the psychics did. It also appeared that the non-psychics compared to the psychics made fewer statements about the health of the target person. Another finding is that non-psychics and psychics made about the same percentage of statements about personality characteristics ("he is dominant"), but that the psychics made much more statements about psychological circumstances ("she has a difficult period"). Finally, it appeared that although the sitter sometimes asked for impressions about circumstances in work, the non-psychics volunteered such impressions (probably relatively even more often than the psychics).

In the set of statements with ppv of the PSY-S series fewer descriptions of psychological characteristics were observed as a logical result of the scoring procedure (WXT: T=1.00, N=12, p<.01). The same pattern was also observed in the PSY-E series and the four control series. Besides this, in the MNP-S and PNP-S series fewer descriptions of physical characteristics were observed in the set of statements with ppv than in the set of statements without ppv (WXT:

MNP-S: T=3.50, N=11, $p<.01$; PNP-S: T=5.00, N=10, $p<.02$), while in the PSY-S series more descriptions of physical characteristics were observed in the set of statements with ppv than in the set of statements without ppv (WXT: T= 4.00, N=12, $p<.01$). This difference between the series with psychics and non-psychics is especially due to the fact that the non-psychics hardly made statements about persons related to the target person (see also the next paragraph) while psychics often concentrate especially on these characteristics when they discuss persons related to the target person. Most of the statements of the non-psychics in this category are statements about information present in the photograph concerning the target person ("she is decently dressed", "he is about 50"), in order to say at least something.

(2) Person discussed in the topics

Out of the 413 topics discussed in the statements of the MNP-S series a total of 9 (2%) concerned a person related to the target person. In the PNP-S series 1 of the 335 topics (0%); in the MNP-E series 6 of the 389 topics (2%) and in the PNP-E series 1 of the 255 topics (0%). These figures were in the PSY-S series 178 of the 1428 topics (12%) and in the PSY-E series 130 of the 1138 topics (11%). The distributions in the series with psychics and non-psychics significantly differ in the standard series as well as in the experimental series (MWU: U varies from U=11.00, $p<.002$ for the PSY-S vs the PNP-S series to U=28.00, $p<.05$ for the PSY-E vs the MNP-E series). Thus, non-psychics hardly volunteer impressions about persons related to the target person.

Further, it appeared in the PSY-S series that, compared to statements without ppv statements with ppv involved relatively the same proportion of topics concerning a person related to the target person.
Since the total number of statements involving impressions about persons related to the target person in each of the control series is about nil it is impossible in the case of the non-psychics to compare the sets of statements with and without ppv with respect to this characteristic.

(3) Number of statements about past, present and future

For the different series the distributions of statements about past, present and future are presented in Table 5.

TABLE 5
Distribution of statements about past, present and future

	past	%	present	%	future	%
PSY-S series	169	16%	792	75%	99	9%
MNP-S series	34	11%	275	89%	1	0%
PNP-S series	29	11%	225	89%	0	0%
PSY-E series	99	12%	619	78%	80	10%
MNP-E series	37	13%	251	87%	0	0%
PNP-E series	37	19%	153	81%	0	0%

The distributions of statements about past, present and future in the PSY-S and PSY-E series on the one hand and in the corresponding series with the groups of non-psychics on the other hand are significantly different, especially because non-psychics appeared not to make statements about the future (chi-square value varies from chi-square=24.79, df=2, $p<.001$ for the PSY-E vs the PNP-E series to chi-square=36.75, df=2, $p<.001$ for the PSY-S vs the MNP-S series).

After splitting up the sessions into first and second halves no difference in number of statements about the past is observed in the PSY-S and PSY-E series (Wilcoxon Matched-Pairs Signed Ranks Test: T=37.00, N=12, n.s. in the PSY-S series and T=18.00, N=10, n.s. in the PSY-E series; all Wilcoxon T tests applied in this paper are two-tailed; table G of Siegel's Nonparametric Statistics, McGraw-Hill, 1956, is consulted to find the significance level of the observed T-value; the Wilcoxon T test is further denoted as WXT). The same holds for all control series (WXT: T varies from T=6.00, N=8, n.s. for the MNP-S series to T=9.50, N=7, n.s. for the MNP-E series).

In the PSY-S and the PSY-E series the set of statements with ppv contained relatively more statements about the past (29% in the PSY-S and 22% in the PSY-E series) and relatively fewer statements about the present and the future compared to the set of statements without ppv (15% in the PSY-S and 12% in the PSY-E series). This difference was also found in the control series. The set of statements with ppv in the MNP-S and PNP-S series combined contained 25% statements about the past, whereas the set of statements without ppv contained 10% statements about the past. In the MNP-E and PNP-E series combined the percentages are 27% and 13 % respectively.

(4) Number of statements about a favourable, neutral or unfavourable state of affairs

The distributions of statements about a favourable, neutral or unfavourable state of affairs are presented in Table 6.

TABLE 6
Distribution of statements concerning a favourable, neutral or unfavourable state of affairs

	favourable	%	neutral	%	unfavourable	%
PSY-S series	209	20%	377	35%	474	45%
MNP-S series	97	31%	156	50%	57	19%
PNP-S series	42	17%	146	57%	66	26%
PSY-E series	110	14%	318	40%	370	46%
MNP-E series	81	28%	155	54%	52	18%
PNP-E series	23	12%	134	71%	33	17%

The distributions in the PSY-S series compared to the MNP-S series and the PNP-S series and the PSY-E series compared to the MNP-E series and the PNP-E series are significantly different (chi-square value varies from chi-square=77.87, df=2, p<.001 for the PSY-E vs the MNP-E series to chi-square=43.30, df=2, p<.001 for the PSY-S vs the PNP-S series). The psychics made more statements about an unfavourable state of affairs than each group of non-psychics (MWU: U varies from U=6.00, p<.002 for the PSY-E vs the MNP-E series to U=35.00, p<.05 for the PSY-S vs the PNP-S series). The percentage of statements about a favourable state of affairs in the MNP-S series and the PNP-S series is not significantly different (MWU: U=47.00, n.s.). The same applies for the MNP-E and PNP-E series (MWU: U=38.00, n.s.).

After splitting up the sessions of the PSY-S series into first and second halves, it appeared that statements about favourable and unfavourable states of affairs were equally distributed over both halves of the session (WXT: e.g. unfavourable: T=32.00, N=12, n.s.). The same pattern was observed in the PSY-E series and each of the series with the non-psychics.

In the PSY-S series the set of statements with ppv had the same distribution with respect to this characteristic as the set of statements without ppv (WXT: e.g. unfavourable: T=28.00, N=12, n.s.). The same holds for the PSY-E series and for each of the four control series.

(5) Number of statements in the form of advice

In the PSY-S series, 60 statements (6%) involved advice. In the PSY-E series this number was 49 statements (6%). As might be expected, no statements of the non-psychics involved advice.

(6) Number of times a silence precedes a statement

In the PSY-S series, 168 statements (16%) were preceded by a silence of 3 seconds or more. In the MNP-S series this number was 101 (33%) and in the PNP-S series, 47 statements (19%). In this respect the psychics differed from the ´matched´ non-psychics (MWU: U=35.00, $p<.05$) but not from the ´professional´ non-psychics (MWU: U=72.00, n.s.). The statements of the ´matched´ non-psychics were more often preceded by a silence.
In the PSY-E series, 90 statements (11%) were preceded by a silence. In the MNP-E series this number was 83 statements (29%) and in the PNP-E series, 52 statements (27%). In the experimental series the statements of both groups of non-psychics were not significantly more often preceded by a silence than the statements of the psychics (MWU: U=36.00, n.s. for the PSY-E vs the MNP-E series and U=38.50, n.s. for the PSY-E and PNP-E series).

Further, it appeared that with respect to this characteristic in the PSY-S series the set of statements with ppv did not significantly differ from the set of statements without ppv (WXT: T=35.50, N=12, n.s.). In the PSY-E series and in each of the four control series no difference was observed between the sets of statements with and without ppv either.

(7) Number of positive and rhetorical statements

Rhetorical statements are statements in which the psychic or non-psychic asks for immediate feedback in contrast with positive statements in which the psychic or non-psychic does not ask for immediate feedback. The distributions of positive and rhetorical statements for the different series are presented in Table 7.

TABLE 7
Distribution of positive and rhetorical statements

	positive	%	rhetorical	%
PSY-S series	737	70%	323	30%
MNP-S series	277	89%	33	11%
PNP-S series	238	94%	16	6%
PSY-E series	559	70%	239	30%
MNP-E series	242	84%	46	16%
PNP-E series	175	92%	15	8%

The distributions of positive and rhetorical statements in the PSY-S series compared to the MNP-S and PNP-S series and the PSY-E compared to the MNP-E and PNP-E series are significantly different (MWU: U varies from U=10.50, $p<.002$ for the PSY-S vs the PNP-S series to U=17.50, $p<.02$ for the PSY-E vs the MNP-E series). The psychics make more rhetorical statements than the non-psychics in both conditions. It has to be noted that especially rhetorical statements tend to elicit feedback from the sitter. Therefore, the lower number of rhetorical statements made by the non-psychics might be a second factor (for the first one see 'the quantitative analysis') that might have contributed to the finding that the sitters in the series with non-psychics provided less feedback responses than the sitters in the series with the psychics.

The difference between the two groups of non-psychics with respect to this characteristic is not significant in the standard series (MWU: U=71.00, n.s.) and in the experimental series (MWU: U=49.00, n.s.).

When first and second halves of the sessions of the PSY-S series

were compared, a significant difference appeared between the two halves of the sessions as regards the percentage of rhetorical statements. More rhetorical statements were found in the first half of the sessions (WXT: T=4.00, N=10, p<.02). The predominance of rhetorical statements in the first half of the sessions was also significant in the PSY-E series (WXT: T=7.00, N=11, p<.02). The difference between the two parts of sessions with psychics was not observed in any of the series with the control groups (WXT: T varies from T=6.00, N=6, n.s. in the PNP-E series to T=5.00, N=5, n.s. in the MNP-S series).
Psychics make rhetorical statements in order to receive immediate feedback. They want to have more information about the target person before giving next impressions, especially in the first part of the session. On the other hand, non-psychics probably feel insecure during the entire session.

The set of statements with ppv in the PSY-S series contained much more rhetorical statements compared to the set of statements without ppv (WXT: T=0.00, N=12, p<.01). This was also found in the PSY-E series (WXT: T= 2.00, N=10, p<.01). However, this difference was not found in any of the series with the non-psychics.
Probably, psychics use rhetorical statements when they verbalize spontaneous and specific impressions in order to reduce the concommitant uncertainty as much as possible. On the other hand, non-psychics seem to feel insecure about all their statements.

(8) Number of statements preceding an informative action

Since all sorts of statements might invite feedback it was studied how many statements the psychic or non-psychic made before the sitter reacted with an informative action. If the psychic made a statement and the sitter reacted directly with one or more informative actions, it is indicated by P1S in Table 8. If the psychic made two statements before the sitter reacted, it is indicated by P2S, etcetera.

The distribution of the PSY-S series is significantly different from the distributions of the MNP-S and PNP-S series and the distribution of the PSY-E series is significantly different from the MNP-E and the PNP-E series (chi-square value varies from chi-square=39.11, df=4, p<.001 for the PSY-S vs the MNP-S series to chi-square=21.31, df=4 p<.001 for the PSY-S vs the PNP-S series). This especially implies that the sitters in the sessions with the non-psychics gave less immediate feedback reactions (P1S) and more delayed feedback reactions

TABLE 8

Distribution of statements preceding an informative action

	statements	P1S%	P1S	P2S	P3S	P4S	P>4S
PSY-S series	1060	45%	482	136	42	16	19
MNP-S series	310	20%	62	25	18	7	14
PNP-S series	254	26%	67	28	8	9	10
PSY-E series	798	49%	390	114	24	9	11
MNP-E series	288	30%	87	26	17	5	9
PNP-E series	190	21%	39	12	13	1	0

(P>1S) compared to the sitters in the sessions with the psychics (MWU: U varies from U=16.50, p<.002 for the PSY-S vs the MNP-S series to U=23.00, p<.02 for the PSY-S vs the PNP-S series).

In the PSY-S series the set of statements with ppv contained an equal proportion of statements which were part of P1S interactions as the set of statements without ppv (WXT: T=16.00, N=12, n.s.). A similar result was found in the PSY-E series and each of the four control series.

The interactional analysis

The objective of this analysis was to study the actions taken by the psychics and non-psychics after receiving a denial to one of their statements. In the case of the standard series with the psychics it was found that they mainly use four types of responses to a denial:

(1) Accepting the denial by giving another impression (AD)
(2) Giving a new interpretation to the denied impression (NI)
(3) Suggesting the target person knows better than the sitter (T>S)
(4) Suggesting the content of the informative action is equal to the content of the statement (I=S)

It was conceivable that other categories of reaction to a denial had to be used in the case of non-psychics. This appeared not to be the case. The non-psychics used the same type of actions after receiving a denial. However, it appeared that the non-psychics never used the reaction which involves the suggestion that the target person knows better than the sitter. In Table 9 the different responses of the psychics for the different series are presented.

TABLE 9
Distribution of types of responses to a denial

	AD	NI	T>S	I=S
PSY-S series	11	50	13	8
MNP-S series	22	9	0	9
PNP-S series	10	3	0	2
PSY-E series	18	53	15	24
MNP-E series	32	13	0	6
PNP-E series	8	4	0	1

Psychics most often responded with a new interpretation to the denied content of the impression, while non-psychics most often accepted the denial by giving another impression (Neglecting T>S and I=S responses: MWU: U varies from U=6.00, $p< .002$ for the PSY-S vs the MNP-S series to U=24.00, $p<.02$ for the PSY-E vs the PNP-E series). It further appeared that the ´matched´ non-psychics more often suggested that the content of their statement was equal to the informative action than the ´professional´ non-psychics did. Although the difference between the figures in each series is not significant, the difference might be a third indication (see also ´the informational analysis´) that the ´matched´ non-psychics took their impressions more serious (believed in the potential correspondence between the content of their statements and the facts concerning the target person) than the ´professional´ non-psychics did.

DISCUSSION

As regards the length of the sessions, it seems safe to conclude that, despite the differences in procedure (asking for paranormal

impressions versus asking for ´uninhibited´ impressions) and in the behaviour of the sitters (not asking questions about certain topics versus asking questions about certain topics; giving more versus less feedback), the number of statements made by psychics in a session is much higher than the number of statements made by non-psychics. In contrast to non-psychics psychics have a lot of experience with such a task.
As regards the two control groups, it appears that the ´matched´ non-psychics (persons matched in age, sex and education with the psychics) made more statements than ´professional´ non-psychics (doctors, psychologists and laywers), in the case that only an object belonging to the target person was presented. The ´professional´ non-psychics who all have a higher level of education compared to the ´matched´ non-psychics, probably had still more doubt about the feasibility of the task.

In the standard and experimental series with psychics it was found that approximately only one out of ten statements could be considered as meeting the criterion of ´being sufficiently spontaneous and specific´, i.e. could be assigned potential paranormal value. Also, that approximately only one out of ten of the statements with potential paranormal value could be considered to meet the criterion of ´being sufficiently correct´, i.e. could be assigned positive paranormal value. According to the author´s estimation none of the statements which met both criteria could be considered to have a probability lower than 1 in 100.
The same number of statements with potential paranormal value and the same number of statements with positive paranormal value were found in the standard and experimental series with each group of non-psychics. Therefore, the conclusions from the previous studies that it is not necessary to assume a specific paranormal ability of psychics, and that research with psychics does not offer a more promising research method than the methods currently employed in parapsychology, are confirmed by the data of this control study.

As regards the behaviour of psychics, ´matched´ non-psychics and ´professional´ non-psychics, the following findings result from the structural analysis:
(1) The two groups of non-psychics volunteered at least as much impressions about circumstances in work of the target person as psychics did. While psychics and non-psychics volunteered the same proportion of impressions about personality characteristics of the target person, only psychics volunteered impressions about

psychological circumstances of the target person. Also, only psychics gave impressions about the physical health of the target person.
(2) In contrast to non-psychics, psychics volunteered impressions about persons related to the target person. The most important reason for this might be that psychics, in direct contact with their clients, try to convince them with respect to their ability especially by giving descriptions of physical characteristics of persons from the the environment of the sitter.
(3) In contrast to psychics, non-psychics made no statements about the future of the target person.
(4) Psychics gave much more impressions about unfavourable states of affairs concerning the target person than non-psychics. These impressions especially concern the physical health and psychological circumstances of the target person.
(5) In contrast to psychics, non-psychics made no statements involving advice to the target person.
(6) Statements from the ´matched´ non-psychics in the standard series were more often preceded by a silence than statements from psychics and from the ´professional´ non-psychics. This might reflect a difference in experience in dealing with unknown people between psychics and ´professional´ non-psychics on the one hand and ´matched´ non-psychics on the other hand.
(7) Compared to statements from non-psychics, statements from psychics more often had a rhetorical form, indicating that psychics expected direct feedback from the sitter.
The higher number of rhetorical statements from the psychics might be one of the reasons why the the sitters in the sessions with psychics gave more feedback than the sitters in the sessions with non-psychics. Another reason might be that the sitter in the sessions with the non-psychics asked for impressions about certain topics in the case the subject did not volunteer such impressions. This active behaviour (asking questions) might to some extent have suppressed the reactive bahaviour (giving feedback) of the sitter.

Fewer descriptions of psychological characteristics were observed in the set of statements with ppv compared to the set of statements without ppv (as a logical consequence of the scoring procedure). This applied to both psychics as well as non-psychics.
Another difference between the set of statements with ppv and the set of statements without ppv of the standard and experimental series with the psychics was that a much higher number rhetorical statements was found in the sets of statements with ppv. In the discussion of the standard series with psychics (see Boerenkamp 1985b) it was suggested

that the psychic more often invites immediate feedback when he verbalizes a spontaneous or specific impression in order to reduce as much as possible the concommitant uncertainty. This relationship between potential paranormal value and rhetorical form of statements was not found in any of the series with non-psychics. The non-psychics probably felt insecure about all their statements.

The data of the interactional analysis in the standard and experimental series indicated that the psychics most often reacted to a denial by giving a new interpretation to their impression, while the non-psychics most often reacted to a denial by accepting it. Further, in this respect the ´matched´ and ´professional´ non-psychics seemed to differ. Although statistically not significant, the ´matched´ non-psychics compared to the ´professional´ non-psychics more often suggested that the content of their statement was equal to the informative action. This difference between the two groups of non-psychics was interpreted as one of the indications that the ´matched´ non-psychics took their impressions more seriously (as maybe really applying to the target person) than the ´professional´ non-psychics did. Two other findings might indicate that ´matched´ non-psychics in this respect resembled more the psychics than the ´professional´ non-psychics. The first one is the higher number of statements from the ´matched´ non-psychics in the case that only an object belonging to the target person was presented. The second one is the finding that two ´professional´ non-psychics made some very specific statements after having made an explicit reference to psychic activity ("I am not a clairvoyant but I will try to be one if you wish"), expressing their unbelief in the feasibility of the task. All ´matched´ non-psychics seemed to take the task seriously.

ABSTRACT

The aim of the analyses presented in this paper was to compare ´paranormal impressions´ of 12 psychics with those of two groups of 12 non-psychics. The statements reflected impressions about persons unknown to the subjects. The groups of non-psychics were (1) a group of randomly selected persons, matched in age, sex and level of education with the psychics (´matched´ non-psychics) and (2) a group of doctors, psychologists and laywers who, because of their profession, were equally experienced to the psychics in dealing with problems of people (´professional´ non-psychics). The sessions

involved (a) one standard condition in which the subject was given a photograph and an object of the target person and received feedback from the sitter and (b) one experimental condition in which the subject was given only an object of the target person and received feedback from the sitter.

It appeared that the group of psychics made a much higher number of statements than each group of non-psychics and that the ´matched´ non-psychics made a higher number of statements than the ´professional´ non-psychics in the case that only an object belonging to the target person was presented.
Each of the three groups produced in the two conditions about the same percentage of statements with positive paranormal value, i.e. spontaneous and specific statements which are correct. Only about 1 percent of all statements are spontaneous, specific and correct (applying to the target person) but that can be expected by the chance-hypothesis. The data of the present study confirm the conclusion from previous studies, i.e., that research with psychics does not offer a more promising research method than the methods employed usually in parapsychology, using non-psychics as subjects. As might be expected, the structure of the behaviour of non-psychics is considerably different from the behaviour of psychics. Among other things: In contrast to non-psychics, psychics especially volunteer impressions about the physical health and psychological circumstances of the target person, about persons related to the target person, about future events and about unfavourable states of affairs. Some of the statements of psychics involve advice to the target person. Statements of psychics have more often a rhetorical form.
Psychics mosttimes react to the denial of a statement with giving a new interpretation to the denied impression, while non-psychics mosttimes react to a denial with acceptance of the denial.

REFERENCES

Boerenkamp, H.G., Schouten, S.A. ´Estimating the potential paranormal value of verbal statements, J.o.P., 47, 1983, 121-130.

Boerenkamp, H.G. ´Potential paranormal value of statements of psychics acquired under feedback conditions, E.J.P., 5, 2, 1984, 101-124.

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part I. Experimental design´, E.J.P., 5, 4, 1985a, 327-348.

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part II. The standard series´, E.J.P., 5, 4, 1985b, 349-371.

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part III. The first group of experimental series´, E.J.P., 6, 1, 1985c, 33-70.

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part IV. The second group of experimental series´, E.J.P., 6, 2, 1986, 107-128.

ACKNOWLEDGMENT

The author gratefully acknowledges his indebtedness to the assistant-researchers L. Heijdendael and R. Flohill for carrying out the sessions with the control groups.

EUR.J.PARAPSYCHOL., V.6, 1986

EVALUATING FREE-RESPONSE RATING DATA

Sybo A. Schouten
Gert Camfferman
University of Utrecht

During the last decades the use of forced-choice methods in experimental research in parapsychology has gradually declined in favor of free-response techniques. In free-response studies complex targets like pictures are employed and the percipient´s mentation during the session is compared with the content of the target for that trial. To enable statistical assessment the mentation is matched also either with the targets for the other trials or with a number of controls, other potential targets from the pool from which the actual target was randomly selected. In the latter case the assessment is based on the difference in agreement of content between mentation and target and between mentation and controls. Free-response techniques are more representative of paranormal phenomena as they occur in daily life. They are generally be more interesting to the participants and have the advantage of allowing process analysis in the case of significant ESP results. The content of the percipient´s mentation during the session can be studied in relation to the content of the target picture.

An obvious disadvantage of free-response techniques is that they are rather time consuming. One session which might take up to one hour yields in general only one trial. Thus free-response studies involve most often a rather limited number of trials. This disadvantage is aggravated by the fact that free-response studies are more difficult to analyze statistically than forced-choice studies. Most statistical techniques applied in free-response studies are actually adaptations from techniques developed for forced-choice methods. As a consequence the free-response data are more or less reduced to forced-choice data resulting in a loss of sensitivity. When for instance forced matchings are applied, in which the set of targets are paired with the set of responses, a time consuming free-response study results statistically in comparing the number of correctly matched pairs of targets and responses with the number of incorrectly matched pairs. In other words, if the study involved 25 trials which perhaps took 25 hours to collect, it has become equivalent to one run of forced-choice card matching which might have been obtained in less than 5 minutes. This discrepancy in time investment between free-response and forced-choice studies seems only acceptable if it can be proven that either free-response studies are more sensitive for detecting ESP or that knowledge is gained from the process analysis which free-response studies allow. These two potential advantages of free-response studies require, however, more sensitive techniques for analysing free-response data than evaluations based on hit/miss ratios which are used with forced-choice methods.

The existing techniques can be classified into ranking and rating techniques. Ranking techniques involve methods in which for each response (or target) the targets (or responses) are ranked in order of agreement of content. In the simplest case, that of forced matchings, all targets and responses of the study are compared in such a way, that each response is paired to the target which is most similar in content under the condition that that response-target pair is withdrawn from the pool of responses and targets. Thus each response is paired with a different target. Burdick and Kelly (1977) discuss methods for evaluating data based on forced matchings.

A more sensitive method seems preferential ranking. In preferential ranking a judge ranks all targets (or responses) used in the trials of the experiment against each response (or target) in order of agreement between content of response and targets. However, evaluation methods for preferential ranking assume independence between the ranking orders of the targets for each of the responses. As various authors have noted, most often this assumption will be violated. According to Kennedy (1979) the preferential ranking method can be most misleading under the very conditions needed to get highly significant results. He suggests a change in judging procedures to ensure independence between the rankings for the various responses; for instance, by the inclusion of dummy targets. Thalbourne (1979) provides another solution. Especially in the case of few trials, which is rather common for free-response studies, a randomization test can be applied which yields an exact probability for the rankings of the actual targets observed.

Solfvin, Kelly and Burdick (1978) provide solutions for cases of preferential rankings whereby weights or scores are assigned to each rank to increase the sensitivity of the method. They also assume independence between the rankings of the different responses. Statistical methods are presented and tables of exact probabilities are provided for those cases in which non-normality of the data have to be assumed.

Another method often employed in free-response studies is one that employes different target sets for each trial and to have the subject assign ratings to all pictures of the set. A target set consists of a number of pictures from which randomly one is selected to serve as the actual target in the experiment. The others are used as controls. The rating values assigned to pictures are based on the agreement between mentation (reported or not) and the content (or perhaps symbolic meaning of the content) of the pictures. Based on the ratings assigned for each response, the pictures can be ranked and one of the familiar evaluation methods for preferential ranking may be applied. Independence of ratings for the different trials is easily ensured. But by turning ratings into ranks the greater sensitivity which the rating method might yield is lost. Hence a statistical evaluation is needed which does credit to the higher sensitivity which ratings might offer. To this end most often z scores are applied, first used and reported by Stanford and Mayer (1974).

Stanford z score method for analysing rating data

With this method for each trial a z score, which we will call Stanford z score to discriminate it from other standard normal scores, is computed by subtracting the mean of all ratings from the rating assigned to the actual target and to divide this difference by the standard deviation of the ratings from which the mean was calculated. In Stanford and Sargent (1983) properties of this z score are discussed and it is stated (p.320) that such z scores "can be used as the dependent variable in ESP studies whether that dependent variable is being correlated with some other variable (e.g., extraversion), is being treated with a complex parametric statistic (e.g., the ANOVA), or is being used as the basis for infering the occurrence of ESP with a given group or condition." However, it is noted that the shape and variance of the obtained distributions of such scores will not necessarily be normal depending on how subjects assign ratings but it is added that the problem here is "neither more nor less troublesome than in other cases in the behavioral sciences." In the case of considerable non-normality nonparametric tests are advised, among others because in such cases nonparametric tests can be more powerful than parametric ones.

When analysing free-response rating data of an experiment applying nonparametric tests on the Stanford z score distribution of the targets a significant result was observed indicating psi-missing. However, it became soon clear that the result was purely artifactual and could be explained by the rating behavior of the subjects. This led us to study the properties of the Stanford z scores in more detail. From the findings, which we describe in the following paragraph, we conclude that one should be rather careful when basing the analysis on Stanford z scores regardless of whether parametric or nonparametric tests are used.

Hansen (1986) reported that the Stanford z score distribution of ratings on a 40-point scale in the case of a target pool of 4 and under the condition that no ties are allowed, e.g., that all ratings assigned in a trial differ from each other, yields a bimodal distribution. We studied z score distributions for a number of conditions. The distributions were obtained by generating all possible permutations of assigning ratings for different types of rating behavior or by taking a randomly selected sample from these permutations. Rating behavior will most likely influence the use of extreme values, for instance, by assigning more values of 0 or 30 (on a 30-point scale). Therefore, distributions were generated like for instance 0 x x x 30, which represents rating behavior in which subjects assign one zero rating and one rating of 30 but select the other three ratings from the range 0-30 inclusive with equal probability for each value. Figure 1 presents a number of distributions obtained for various types of rating behavior.

From figure 1 it appears that the distributions are in all cases non normal and only symmetrical when subjects select ratings with equal probability from the whole range. Decreasing the size of the target set yields flatter distributions. Decreasing the range of the ratings results in more irregular distributions. All distributions have an upper and lower limit of z scores. In cases where subjects select ratings with unequal probabilities from the range applied the distributions become asymmetrical. We suspect

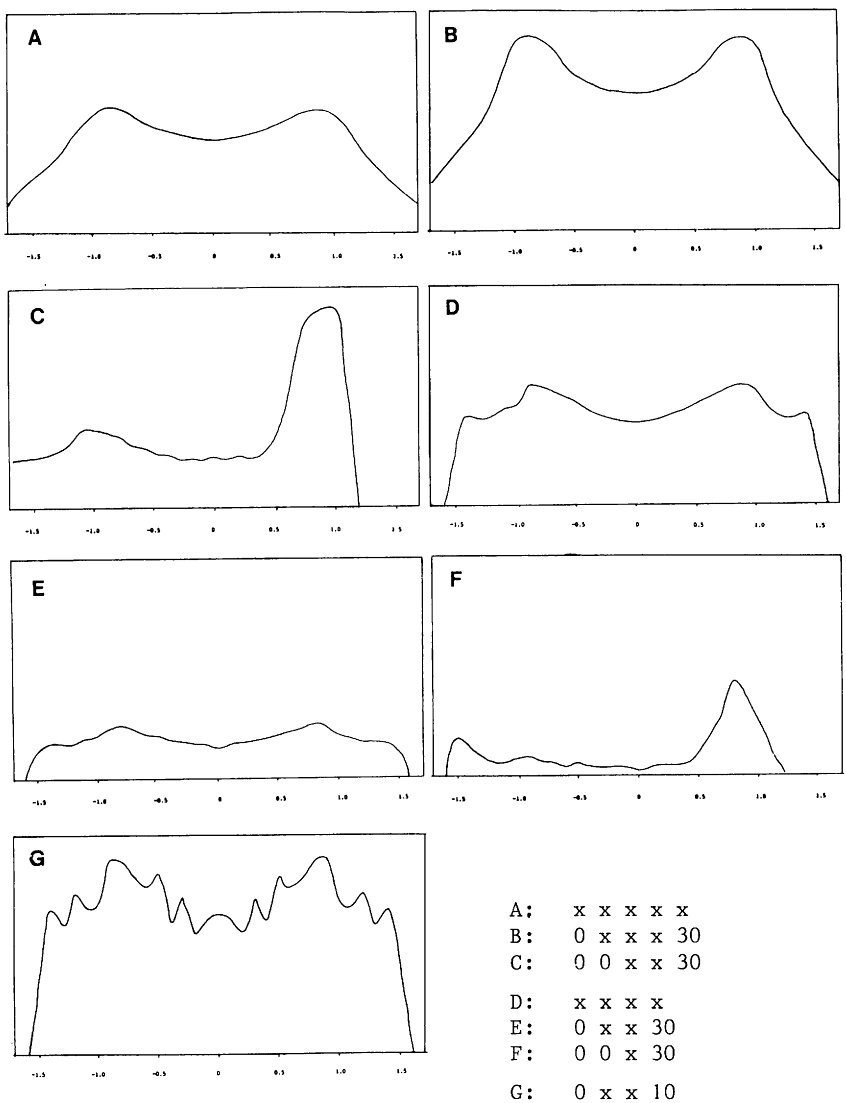

FIGURE 1
Distribution of Stanford z scores for various types of rating behavior
Scale from z = -1.50 to z = 1.50

that as a rule such asymmetrical distributions will be found since a number of subjects will have the tendency to assign one value greater than zero to the picture which is judged most similar to their mentation, while the other pictures are then assigned ratings of zero. Hence it can be concluded that rating behavior influences the distributions of Stanford z scores.

Stanford z scores are also peculiar in some other respects. Their value and range is rather sensitive to the number of equal ratings assigned. That is of relevance since often a number of ratings of zero will be assigned. In the case equal values are assigned the actual size of the ratings have no influence any more on the size of the Stanford z score. For instance, the rating patterns 0-0-1-1-1 and 1-1-0-0-0 yield z scores of -1.10 or + 1.10 respectively for the target, assuming the target to be positioned first, and these z values remain the same for all cases in which the first two and last three ratings are equal. Thus 1-1-0-0-0 yields the same Stanford z score for the target as 100-100-0-0-0. The same applies to comparable situations like 1-0-0-0-0 and 100-0-0-0-0; in both cases the target receives a Stanford z score of +1.72. Hence the Stanford z scores does not always reflect the similarities or differences between mentation and targets which subjects express in their assignment of rating values.

A third complication is that when relatively many ratings of equal value are assigned the z score distribution tends to become discrete rather than continuous. Especially since free-response studies in general involve few trials the discrete character of such distributions violate assumptions on which many parametric and nonparametric tests are based. This might result in an incorrect statistical evaluation.

These properties of the Stanford z scores suggest that the use of parametric tests to determine ESP from z score data might in many cases be disputable. This was already observed by Stanford and Sargent (1983, p.321) although they state that many parametric statistics are known to be quite robust against type I statistical errors even when the shapes of the distributions vary considerably from normal. However, considering the on average small number of trials of free-response studies we feel that this argument should be handled with care.

A more important problem seems to us that in many cases the conditions of the experiment will influence the rating behavior of subjects. In most studies the aim is to evaluate the effect of an independent variable on the dependent one, the latter expressed in z scores. In that case the main analysis is focussed on possible differences in ESP scoring between conditions. That implies that an influence of conditions on the rating behavior, and consequently on the z scores, must be eliminated before a proper evaluation of the difference as regards ESP scoring can be made. Differences in rating behavior due to conditions might occur under various circumstances. For instance, comparing Ganzfeld versus non-Ganzfeld might result in different and perhaps richer mentation in the Ganzfeld condition and therefore to the assignment of on average higher ratings in that condition. That might, but not necessarily so, contribute to a difference in z score distributions for the two conditions.

In a free-response study (Schouten 1987) a marginal significant difference in distribution of ratings was observed between two conditions which differed as regards familiarity of the targets to the percipient. A possible effect of target familiarity on rating behavior was already discussed by Irwin (1982). A similar difference in rating distributions can be expected in the case targets differ in amount of details. Such differences will often be reflected in the number of zero ratings the subject assigns. Hence z score distributions from different conditions might be compared which for instance strongly differ in degree of skewness. Given the often discontinuous character of the data this implies that care should be taken even when nonparametric tests are applied. For instance, in the case that two conditions are compared, one in which on average four zero ratings are assigned whereas in the other condition on average three zero ratings are assigned, with a nonparametric test like the sign test a significant difference only due to rating behavior can already be found for an n as low as n=11. Admittedly, this is a rather extreme example but it might serve as an illustration that it is not true that in the case parametric tests are not applicable nonparametric tests will always do.

Hence, the main disadvantages of using Stanford z scores seem to us that different rating behavior in conditions might have an effect on the evaluation of differences in ESP scoring between the conditions; that in many instances differences in degree of similarity between targets and mentation is not reflected in the Stanford z scores; and that the distributions are in general not normal, asymmetric and not continuous. To meet these objections a different evaluation procedure is proposed based on a randomization test. With this test differences in scoring between conditions can be evaluated in such a way that it can be assumed that possible effects of condition on rating behavior are eliminated. In addition it is possible to obtain ESP scores for each trial which compared to Stanford's z scores better reflect differences in degree of similarity between mentation and target or controls, which are more continuous and which can be assumed to be sufficiently normally distributed to allow for instance in the case of related samples to compute a product-moment correlation coefficient between conditions.

A randomization test for rating data

The randomization test is based on the sum of ratings over the trials. Subjects assign ratings to the pictures. Of these pictures one is the target picture actually used in the experiment, the others are control pictures. In the case of assigning rating values to the control pictures it can be assumed that ESP can have no effect on these ratings. If we select randomly from each trial a rating value of a control picture and take the sum of these ratings then based on all possible combinations of ratings for control pictures over the trials a distribution is obtained which will tend to be normal even in the case that the ratings themselves were selected with unequal probabilities. The randomization test provides an answer to the question to what extent the sum of the ratings over the trials assigned to the target pictures deviate from the mean sum of the ratings assigned to the control pictures. Consequently the sum of the ratings assigned to the target pictures is expressed as a standard normal score based on the distribution of the sum of the ratings assigned to the controls. This standard normal score will be called the standardized sum of ratings score or SSR score.

The number of combinations on which the distribution of the sum of ratings for the controls is based equals the number of controls to the power of number of trials. A good approximation of this distribution is obtained by calculating mean and standard deviation from the mean and variance of the ratings for the controls of the individual trials. The mean of the distribution for sum of scores will be equal to the sum of the means of controls for the trials. The standard deviation is found by taking the square root from the sum of the variances for control ratings over the trials.

With this method we contrast the sum of the ratings assigned to the targets with the mean sum of ratings assigned to the controls. It can be shown that in the case the distribution of the sum of ratings is based on the ratings assigned to the controls as well as on the ratings assigned to the target the method of evaluation becomes very insensitive. It appears that in the latter case only the most optimal assigning of ratings, viz. 30 to the target and 0 to each of the controls, yields an SSR score with an associated probability equivalent to what would be obtained if the experiment was treated as a guessing experiment and the number of hits (target picture achieving the first rank) evaluated by the binomial. In all other cases when non-zero ratings are assigned to the controls the binomial based on rankings is more sensitive than the randomization test. Therefore the distribution of the sum of ratings should be based on the sum of ratings for controls only.

Since SSR scores can be assumed to be standard normal its associated probability can be obtained from the standard normal distribution. SSR scores of different condition can be compared because SSR scores can not only be considered standard normal but also independent of differences between conditions in range of ratings, rating behavior or number of controls applied.

ESP scores for individual trials

To obtain ESP scores for individual trials the rating value assigned to the target is converted into a standardized average rating score for the target (SAR score). The conversion is based on the following considerations. It can be assumed that in general subjects do not select all rating values with equal probability for each value but that the condition determines these probabilities. Thus a condition with unfamiliar targets and a short duration for the session so that few spontaneous mentation is obtained might result in the assignment of a relatively high number of zero ratings. On the other hand a condition with complex targets and rich mentation might yield most ratings of non-zero value. Such influence of condition will on the average effect the rating behavior of all subjects and is reflected in the ratings assigned to the controls. Therefore it seems proper to express all ratings of a condition in the distribution of average ratings of controls for each trial. A second reason for this conversion is that it seems attractive to convert ratings into values which can be assumed to be normally distributed. The third reason is that like in the case of the SSR scores this conversion makes each score for the individual trial independent of effects of condition, number of control pictures applied, or differences in range of ratings.

The distribution of the sum of ratings for the controls can be considered as the distribution of ratings associated with that condition. Reduced to the level of individual trials we assume this distribution to be typical for the condition and express all ratings in this distribution of average ratings. Thus all ratings are converted into standard normal scores by computing its distance from the mean of average ratings for the controls of the trials and deviding it by the standard deviation observed for these average ratings.

Then for each trial a standardized average rating or SAR score for the target is defined as the difference between this standard normal score for the target and the average standard normal score for target and controls. Since the SAR scores are based on true standard normal scores, which means scores obtained from a normal distribution, SAR scores can be considered normal too. For each trial the sum of SAR scores for controls and targets is zero. Therefore in the case of related samples we might compare individual achievement over conditions by calculating a product moment correlation between the SAR scores of the two conditions.

As an example table 1 presents rating values, distributions of sum of ratings of controls over the trials, SSR and SAR scores for the data obtained in the two conditions of the experiment described in Schouten 1987. The mean and sd of the distribution of average ratings per trial are derived from the mean and sd of the distribution of the sum of ratings for controls over the trials by dividing these values by respectively n or Vn.

No psi was observed in this study but rating behavior differed for the two conditions. Mean and standard deviation of sum of rating values for controls over the trials were established in two different ways. One method was to take 100,000 samples of sums of rating values over the trials when from each trials one rating value assigned to a control picture was randomly selected. With the other method mean and standard deviation are based on the average value of ratings assigned to controls and associated variance of each trial. The standard deviation of the distribution of sum of ratings for controls is obtained by taking the square root of the sum of the variances of each trial. It turned out that the two methods yielded nearly equal values for mean and sd of the distribution of sum of ratings for controls.

As an example we show how the SAR score for the first trial of condition I was obtained. The mean of the distribution of average

TABLE 1
Ratings, SSR and SAR scores for two free-response conditions

Condition I Ratings					Condition I SAR	Condition I Stanford	Condition II Ratings					Condition II SAR	Condition II Stanford
0	0	0	0	30	-.64	-.45	0	0	15	30	30	-2.47	-1.00
0	0	0	0	10	-.21	-.45	0	0	0	0	0	0	0
0	0	0	0	20	-.42	-.45	0	0	0	0	2	-.07	-.45
0	0	0	0	5	-.11	-.45	0	0	0	0	0	0	0
0	0	0	10	20	-.63	-.67	0	0	0	0	15	-.49	-.45
0	0	0	0	15	-.32	-.45	0	0	0	10	10	-.66	-.73
0	0	0	30	30	-1.27	-.73	20	0	0	10	15	1.81	1.23
15	0	10	20	30	0	0	0	0	0	0	20	-.66	-.45
0	0	0	0	5	-.11	-.45	0	0	0	0	0	0	0
15	0	0	0	3	.63	.45	0	0	3	20	25	-1.58	-.80
0	0	0	0	5	-.11	-.45	0	0	10	20	20	-1.65	-1.00
20	5	5	25	30	.32	.26	5	0	15	30	30	-1.81	-.79
0	0	0	5	30	-.74	-.54	0	0	0	0	0	0	0
0	0	25	28	30	-1.75	-1.09	0	0	0	0	5	-.16	-.45
25	0	0	0	20	1.69	1.29	15	10	20	25	30	-.82	-.63
0	0	0	10	25	-.74	-.64	0	0	0	0	0	0	0
30	10	20	30	30	.63	.67	25	0	0	0	0	3.29	1.79
10	0	20	25	30	-.74	-.58	0	0	2	10	10	-.72	-.85
9	0	0	0	25	.23	.20	0	0	0	0	0	0	0
0	0	0	0	5	-.11	-.45	0	0	0	0	0	0	0
20	5	5	5	20	.95	1.10	20	0	0	5	10	2.14	1.55

Distribution of controls M: 184.5 sd: 43 SSR score=-.94 | M: 114 sd: 28 SSR score=-1.04

Product-moment correlation between SAR scores: r=.07

ratings per trial for this condition equals 184.5/21=8.786, the sd equals 43/V21=9.383. Expressed in this distribution the rating values become -.934,-.934,-.934,-.934,2.261. These values have a mean of zero and a sd of 1.47. Then the SAR score for the value associated with the target becomes -.934/1.47=-.64

Properties of the randomization test

Although the randomization test described above seems statistically sound we further studied its properties especially regarding its sensitivity to detect ESP. To this end we simulated by computer an experiment consisting of 20 trials and 5 pictures for each trial, one target and 4 controls, by randomly generating 20 rows of 5 numbers between 0 and 30 inclusive. The first number of each trial is considered the rating value assigned to the target. The effect of two variables on the SSR scores was investigated. One is the rating behavior of subjects, the other the amount of ESP. For rating behavior we manipulated the probability of selecting rating values of zero. In table 2 in which the results of these analyses are presented a rating behavior of 0 indicates that all rating values were chosen with equal probability. Rating behavior of .6 indicates that for each trial the probability of selecting a rating value of zero was .6. In other words, in each trial on average 3 ratings of 0 were assigned. The amount of ESP was operationalized as the number of subjects assigning the highest rating value to the target in addition to what could be expected by chance. Amount of ESP varied between 0 and 5. In table 2 an amount of ESP of 0 means no ESP and hence we can expect on average in 20/5=4 trials or rows the highest rating assigned to the target. An amount of ESP of 2 means that in 2 trials the highest rating value was assigned to the target (by the computer placed in the first position) while for the other 18 trials or rows the random process decided the position of the highest rating. Hence in this case the mean expectation of the highest rating value assigned to the target becomes 2+18/4=5.6.

Table 2 presents the distributions of SSR scores based on the randomization test for each of the combinations of the two variables studied. Each cell is based on 100 simulated experiments. From these data it can be concluded that in most conditions the sensitivity of the randomization test is rather low and less than for instance when a simple binomial test was applied to the number of hits obtained. Only in extreme cases of rating behavior and amount of ESP influence do the SSR scores become more sensitive compared to the binomial test. In all other cases the binomial is preferable. For instance, in the case of ESP=5 when 5+15/5=8 hits can be expected, the binomial yields an exact one-tailed probability of p=.01 whereas the SSR score yields on average a z=1.7 with an associated one-tailed probability of p=.045. A similar simulation procedure was run for experiments in

TABLE 2
SSR score distributions as a function of rating behavior and influence of ESP

ESP influence		Rating behavior 0	.2	.4	.6	.8
0	mean SSR	0	0	0	0	0
	s.d. SSR	1.1	1.2	.9	1.1	1.1
1	mean SSR	.3	.4	.3	.4	.7
	s.d. SSR	.9	1.0	1.0	1.2	1.1
2	mean SSR	.6	.7	1.0	1.1	1.2
	s.d. SSR	.8	1.0	1.0	1.3	1.4
3	mean SSR	1.0	1.0	1.2	1.5	1.9
	s.d. SSR	.9	1.2	1.0	1.2	1.4
4	mean SSR	1.3	1.5	1.6	2.0	2.5
	s.d. SSR	1.0	1.0	1.3	1.2	1.5
5	mean SSR	1.7	1.9	2.0	2.2	2.7
	s.d. SSR	1.1	1.2	1.0	1.2	1.7

which 7 instead of 5 control pictures were used but this did not result in improvement in sensitivity.

In the same simulation studies Stanford z scores were computed. We know that the distributions for z scores are non-normal but leaving this aside we found that for most cells of table 2 the sensitivity of t test evaluations based on Stanford z scores is comparable to the evaluations based on SSR scores. However, SSR scores appear more sensitive compared to Stanford z scores in cases of strong ESP (ESP>3) and extreme rating behavior (>.5).

From these findings some practical conclusions can be drawn. In general we must assume that the ESP influence on the data is relatively little. Hence unless there is reason to expect a strong ESP influence in the experiment the binomial test can be assumed to be more sensitive than an evaluation based on the rating values. The same applies for experiments in which no extreme rating behavior can be expected. For instance, in an experiment in which an atomistic approach to the judging is followed. In that case we expect in general non-zero rating assigned to all pictures and table 2 shows that in that case the SSR scores are rather insensitive.

The randomization test does provide us with a method which is statistically sound but appears to be not sufficiently sensitive to do justice to the potentials of the free-response procedure. Therefore it seems advisable to use ratings mainly to obtain a rank order and to establish individual's ESP scores. Whether the evaluation of conditions and differences between conditions should be based on a ranking evaluation procedure or on a binomial test on hits and misses remains to be seen. It should be of interest to study also the sensitivity of ranking procedures compared to the binomial test. Without more sensitive methods for evaluation and when no process analysis is carried out as seems the case in most free-response studies reported we fail to see the advantages of the free-response method.

ABSTRACT

Free-response studies are very time consuming compared to forced-choice studies. This extra time investment might be justified for instance when free-response studies can be considered to be more sensitive to ESP or when process analysis yields additional information. Free-response studies are evaluated by ranking or rating procedures. Ranking procedures can be considered adaptations from methods developed for forced-choice procedures and might do no justice to the greater sensitivity free-response studies suggest. To obtain better sensitivity rating procedures are applied.

The statistical evaluation of rating data should reflect the potential sensitivity of the rating procedure. Stanford z scores are not considered optimal for this purpose because its distribution is not normal, becomes different in shape and symmetry when rating behavior of subjects differ over conditions, and tend to be discontinuous. A method for evaluating free-response rating data was developed, based on a randomization

test to evaluate the significance of scoring in conditions and between conditions. With this method differences in rating behavior do not influence the evaluation of differences between conditions. Based on a distribution which is considered 'typical' for the rating behavior of the condition involved ratings are converted into scores which can be considered standard normal. Based on this distribution an ESP score for each trial is obtained which can be considered independent of various properties of the condition like rating behavior and number of control pictures applied.

However, further analyses revealed that the sensitivity of the randomization test and of evaluations based on Stanford z scores is rather low. It is concluded that unless more sensitive evaluation methods are developed an evaluation based on rankings or on a binomial test seems preferable to an evaluation based on ratings.

ACKNOWLEDGMENT

This paper has been presented at the 30th Annual Convention of the Parapsychological Association, Edinburgh, August 5-8, 1987.

LITERATURE

Hansen, G. 'A look at z-scores (standardized ratings)'. Research in Parapsychology 1985, The Scarecrow Press, Metuchen, NJ and London, 1986, p.93.

Irwin, C.L. The role of memory in free-response ESP studies: Is target familiarity reflected in the scores. Journal of the American Society for Psychical Research, 1982, 76, 1-21.

Kennedy, J.E. Methodological problems in free-response ESP experiments. Journal of the American Society for Psychical Research, 1979, 73, 1-13.

Solfvin, G.F., Kelly, E.F., and Burdick, D.S. Some new methods of analysis for preferential-ranking data. Journal of the American Society for Psychical Research, 1978, 72, 93-109.

Schouten, S.A. A test of Heymans theory on paranormal phenomena.

Paper submitted to the 30th Annual Convention of the Parapsychological Association, August 1987.

Stanford, R.G. and Mayer, B. Relaxation as a psi-conducive state: A replication and exploration of parameters. Journal of the American Society for Psychical research, 1974, 68, 182-191.

Stanford, R.G. and Sargent, C.L. Z scores in free-response ESP methodology: Comments on their utility and correction of an error. Journal of the American Society for Psychical Research, 1983, 77, 319-326.

Thalbourne, M.A. A more powerful method of evaluating data from free-response experiments. Journal of the Society for Psychical Research, 1979, 50, 84-107.

LETTERS TO THE EDITOR

From J. Palmer:

I wish to bring to your readers' attention three errors/misprints that appear in the statistical formulas in one of my chapters in our new textbook, Foundations of parapsychology. On p.146, formula 3, the denominator should be S , not S /N . On p.148, formula 5, the term (R-1) in the numerator should be (R+1). On p.151, formula 7, the term I should be T.

John Palmer
Institute for Parapsychology
Box 6847, College Station
Durham, NC 27708
U.S.A.

From D.J. Bierman:

ON THE USE OF EXPERT-SYSTEMS IN PSI-RESEARCH

Demystification of Expert-systems

For many people outside of the field of A.I., so-called Expert-systems have a kind of glamorous appeal and seem to trigger the naive idea that this technology might solve all our problems. The SDI of the US administration for instance is among others based on this

fallacy.
In the article 'Minimizing subject fraud in parapsychological laboratories', (EJP, 6.2) the problem raised is the detection (and prevention) of fraud. It is suggested that Expert-systems could help to solve this problem.

Knowledge based systems (which is the proper label) are systems that differ from traditional Data Base Systems in a few respects. Their knowledge, factual as well as procedural, is represented explicitly (which is thought to be an advantage when it comes to maintenance of these systems). Furthermore these systems are able to infer 'new' knowledge from the present knowledge using well specified inference mechanisms. In Cognitive Science these systems can be used as models for human cognitive functioning (not only Experts but, for instance, also Novices might be modelled in this way, see Jansweyer et al, 1986) although a proper methodology to evaluate the implemented models is still lacking. Generally the systems are capable of using fuzzy knowledge (with uncertainty factors). Heuristic knowledge (rules of thumb) can be inserted which might help to reduce the search space (eg. more probable solutions are tried first).

Implication

The implication of the above description is that

a. The systems quality is heavily dependent upon the quality of the inserted knowledge.

One weak knowledge element might have disastreous effects on the functioning of the total system. Knowledge elicitation, the process of gathering this knowledge before one can insert it into the Expert-system, is an underestimated aspect of the construction of Expert-systems. Global domain knowledge might be extracted from written sources like textbooks. To elicit the more specific, expert or maybe intuitive, knowledge, specialistic procedures like the analysis of thinking aloud protocols of human experts (eg. Breuker and Wielinga, 1987) or the analysis of the results of Learning systems (eg. Bierman and Akkerman, 1986) are required.
It is rather clear that in the domain of magic the otherwise already difficult task of knowledge elicitation is more problematic, if not impossible, especially given the limited resources of the

Parapsychological community.

b. The quality of the systems solution is dependent on the quality of the problem specification

Morris clearly distinguishes 3 protocols, in as far as contact between Magician and (pseudo)Psychic is concerned. These are the verbal protocol, the video protocol and the ´human presence´ protocol. He argues that the last one is preferable although certainly not foolproof. He fails to remark however that (at least for the foreseeable future) Expert-systems will have to rely on the Verbal protocol. That is, they will have to rely on the worst possible problem specification.

CONCLUSION

The reference to Expert-systems as a potential tool to help uncovering Subject Fraud is completely unrealistic.

Morris remarks that another strategy to prevent Subject Fraud to be too harmful, is the use of process oriented experiments with variables that are unknown to the subject. I do not only agree but I would strongly advocate this type of experiments on scientific grounds. It is time that each contribution to the field should state its theoretical relevance. Instead of refusing articles because they do not reach significance, the publication policy should result in refusing articles which fail to indicate this theoretical relevance. Since experiments with theoretical relevance tend to have some relevant variables, this publication policy would stimulate research which is rather unsensitive for Subject Fraud.

REFERENCES

Bierman D.J. & Akkerman, A.E. ´The development of knowledge based psychodiagnostic systems´. In: F.J. Maarse, L.J.M. Mulder, W.P.B. Sjouw, en A.E. Akkerman (Eds.) Computers in psychology: Methods, instrumentation and psychodiagnostics, Lisse, Swets and Zeitlinger, 1987, 73-78.

Breuker, J. and Wielinga, B.J. ´Use of models in the interpretation of verbal data. In: Knowledge Acquisition for Expert Systems, A practical Handbook, New York, Plenum Press, 1987.

Jansweyer W., Elshout, J. & Wielinga, B.J. ´The expertise of novice problem solvers´. In: Proceedings of the 7th European Conference on Artificial Intelligence, Brighton, July 1986, 576-586.

D.J. Bierman
R.I.P.P
Alexanderkade 1
1018 CH Amsterdam
The Netherlands

CONTENTS

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.

EUR.J.PARAPSYCHOL., V.6, 1987

A MULTIVARIATE PK EXPERIMENT. PART I. AN APPROACH COMBINING PHYSICAL AND PSYCHOLOGICAL CONDITIONS OF THE PK PROCESS

Walter von Lucadou
University of Utrecht

Despite the large number of successful PK experiments which has been published during the last decades (see Bauer, v.Lucadou 1979, Honorton 1978) it is obvious that the scientific community is still far from being convinced of the reality of PK phenomena. At first glance this seems to be astonishing since the methodological and experimental methods of parapsychology have improved strongly since Rhine (1943) published his report on the first statistical PK experiments. There has also been developments in theorizing which would allow a reconcilation of psychokinesis with the constructs of modern physics, requiring only some 'weak' modifications (see Schmidt 1985) of physics. The main reason for this lack of conviction seems to be the unreliability of experimental results indicative of the existence of a psychokinetic effect.

However, the fact that we are aware of the unreliability of the results indicates that the usual argument of skeptics (see Alcock 1981) who maintain that in parapsychology only positive results are published cannot be true. A meta-analysis of May, Radin and Thomson (1986) indicates, that 4500 unpublished PK experiments must exist to wash out the overall PK effect in the pooled distribution of 332 published PK experiments (during the period of 1970 to 1985). Under the present conditions the possibility of the existence of such an enormous amount of unpublished PK experiments can be safely excluded.

Nevertheless, the impression may be justified that the unreliability of PK experiments is the only reliable result of this type of research (see Blackmore 1983, Hoevelmann 1984). Usually it is assumed that psychological conditions which cannot be controlled readily are responsible for the observed unreliability. Therefore, in order to control these it seems necessary to know more about these psychological conditions. In PK research, however, few systematic studies on the influence of psychological variables on the PK process have been carried out and the results of these studies do not exhibit any clear structure (for an overview see Rush 1977, Rao 1980).

Most researchers, therefore, expected to find more information about the PK process by studying its ´physical´ properties (see Houtkooper 1983, Schmidt 1973, 1974, Jahn 1981) while trying to optimize the psychological conditions in such a way that a maximal PK effect could be expected (see for instance Stanford´s (1974) recommendations for optimal psi conditions within an experiment). The psychological conditions concerning the (mainly selected) subjects were in most cases only qualitatively described. Since no standardized psychological ´measures´ were used one cannot generalize the findings and comparisons of successful experiments can be made only qualitatively. Another disadvantage of the qualitative description of psychological conditions might be a possible selection error since PK experiments might cause an unconscious post hoc attribution of psychological qualities to the subject by the experimenter due to the outcome of the experiment. With ´unsuccessful´ subjects this unconscious attribution might be quite different from those of successful ones. This might be a problem especially in situations where the psychological conditions are used as ´explanation´ for an unsuccessful experiment!

To solve this problem and also to prevent statistical selection errors it would be necessary to combine quantitative psychological ´measurements´ with (quantitative) measurements on the physical or stochastical process which should serve as PK target. In this way psychological variables could be used as independent variables to predict the outcome of an experiment. This approach involves two problems:

1. One does not know which psychological variables are relevant to the

PK process.

2. One does not know which physical variables may be affected by PK.

From ESP research it is suggested that certain personality characteristics may be relevant (see Eysenck 1967, Palmer 1977, Rao 1980, Johnson and Haraldsson 1984). The sheep-goat construct which plays a dominant role in ESP may serve as an example of such a trait variable. Recently Roll (1985) argued that masculinity may play a similar important role within PK processes. It is plausible that PK might also depend on state variables such as mood, or amount of frustration during the actual PK session. One could also imagine the possible influence of physiological variables. In the experiment reported here, however, only psychological variables were considered.

As regards the physical aspect of the problem usually statistical values are considered as quantitative variables which ´measure´ the anomalous physical process called PK. Especially the hit-rate is often used to ´measure´ the PK success. From a theoretical point of view, however, this is neither a very specific measure nor does it convey sufficient information on physical changes in the underlying physical process. From a statistical point of view the hit rate is a rather poor statistical measure since one can easily imagine situations where even dramatic deviations from randomness (for instance by PK) may leave the expectation value of the hit rate unchanged.

To illustrate the lack of information of a hit rate score about the physical process on which it is based the RNG used by H. Schmidt may serve. With this RNG an impulse produced in a detector by a particle of a radio-active source is used to stop a fast running clock (flip-flop). The final state of this clock represents the random event. Even in the case of a highly significant PK effect (deviation from the expectation value in the hit-rate) one cannot deduce which physical process was affected by PK. It could be that the radio-active source was affected, for instance, by altering the decay rate, or the sensitivity of the detector might have been influenced. Also an influence on the time basis of the clock or on other electronic components could be responsible.

The observational theories (OTs) start with the assumption that the ´psi source´ affects the probability of the collapse of the wave vector which describes the physical state of the system. This assumption, however, is a very general (and thus a weak) one since it is not specified according to which observable (or subspace) the collapse takes place. It would be an important step forward if one could find a specific process, which is affected by PK in a specific way. I call such a process a tracer. Finding a tracer would allow to filter out psi from all the other processes which normally may occur in a piece of matter under consideration (the PK target). The main purpose of the experiment which will be described here was to find such a tracer to identify PK at its origin.

This question is also very important in relation to the assumed goal orientedness of PK (see Schmidt 1975, 1985). If one assumes a general goal orientedness as in Schmidt's model the outcome of any statistical experiment would not depend on physical but only on psychological variables of the observer, assuming that real chance events are used as targets. Schmidt calls this complexity independence. It is important to stress that goal orientedness and feedback mechanism are not identical and can lead to different experimental results (as it was shown by the author in a previous paper (Lucadou 1979) and also recently by Vassy (1985) and Varvoglis (1985) (also see Walker 1977)). A feedback mechanism could of course be used to insert some goal orientedness to a certain process, but this would still imply that the outcome must depend on the properties of the underlying physical process, such as circuit complexity or other physical properties. From the point of view of Schmidt's model one would not expect different outcomes of an experiment with different types of random sources if the psychological and feedback conditions are identical. This is the main conceptual difference between Walker's (Walker 1975) and Schmidt's model. With Walker's model and also within the system theoretical model of Lucadou and Kornwachs (1982, Kornwachs and Lucadou 1977) it is possible that the PK effect also depends on physical variables of the target process. This question was of fundamental importance for the design of the present experiment and is called the generator hypothesis because two different RNG's were compared in relation to their PK sensitivity.

2. HOW TO OPERATIONALIZE PK IN THE CONTEXT OF A MULTIVARIATE EXPERIMENT

In order to avoid the problems of selection and attribution it is necessary to operationalize PK not only in terms of an anomalous behaviour of a physical process, but also in terms of psychological and environmental variables. For this reason PK here will be defined as a (significant) correlation of psychological variables with physical ones which are measured independently during the experimental session. This definition is not in contrast with the usual one which includes the concept of ´influence´. However, it is much more general because it could also describe correlations which are not based on influences. In a feedback situation, obviously, one cannot measure the independent psychological variables simultaneously. Thus all relevant psychological variables were measured before the experimental session. The psychological variables are operationalized by the use of a set of questionnaires and the physical variables by a set of different statistical observables which will be described in detail in paragraph 4.5. Furthermore psychological variables are considered to be reasonably stable.

It is obvious that the notion of correlation is more general than the notion of influence. From the point of view of the statistical evaluation method a correlation does not determine which variable has to be considered as independent and which one as dependent. This distinction could only be inferred from the time order of the measurement of the variables. The psychological variables can be considered as the independent ones since they were measured before the physical ones. In the case of PK, however, this distinction might become questionable since at least in the OTs it is assumed that PK is time independent. Also the possibility of precognition has to be discussed as we will see in part III.

Moreover the notion of correlation would not even require a causal link between the two variables. As an example a partial correlation may serve, or the famous EPR-correlation in quantum physics. It is important to recognize that the definition of PK by means of a correlation between psychological and physical variables is more general and still open to different underlying models than the usual definition in terms of a dependency (see Lucadou 1984). Furthermore, it seems to be the only useful definition within a multivariate experimental design.

Special attention has to be devoted to the problem of possible artificial correlations. In general, one can distinguish two different sources for such artifacts:

1. Statistical artifacts due to improper parametric assumptions of the correlation coefficient applied or the occurrence of accidental

correlations.

2. Technical artifacts due to malfunctions of the apparatus or improper shielding against physiological influences.

Such potential technical artifacts are good examples of partial correlations. If for instance the RNG would produce more hits in the morning than in the evening this might produce a spurious correlation with personality variables, simply because subjects of a specific type of personality may prefer to do the experiment in the morning. It is essential to design the experiment in such a way that such artifacts can be ruled out. The details will be discussed in chapter 6.

3. THE MULTIVARIATE APPROACH AND BRUNSWICK'S LENS MODEL

Parapsychologists are well aware of the unreliability of their results. Most of them (and also their critics) usually compare this with the situation in physics were, fortunately, highly significant results are usually also very reliable. However, in most other empirical sciences this is not the case. Especially in (personality-) psychology the unreliability of experimental results is a familiar methodological problem (see Mischel 1968, Rorer and Widiger 1983). One of the aims of personality psychology is to predict the behaviour of a person in a given situation. For instance, industry companies use personality questionnaires in order to select employees in an optimal way. It is known, however, that personality characteristics gained in this way are very poor predictors for actual human behaviour in specific situations. One even gets the impression that the main progress in personality psychology has been to learn not to overestimate the predictive power of such methods. Technically speaking this means that the correlation coefficients between personality variables and behaviour variables are very small. In the last decades however, methods were developed to enhance the correlation between independent and dependent variables. These methods are called 'multivariate' since they describe both the independent and the dependent part of the (experimental) situation not by a single variable while all the other variables are kept constant, but by a set of variables which should be as complete as possible. With this method it would not be necessary to 'prepare' the system as it is done in physics, but on the contrary, one can allow as much variance as possible for the distribution of the variables involved. These variables may depend upon each other or not. The methods allow to organize (aggregate) these variables in a hierarchical structure which results in classifying concepts or constructs. This method was for instance used by Eysenck for the classification of personality traits (see Eysenck 1953). Wittmann and Schmidt (1983) recently argued that it would also be useful to apply this hierarchical method to organize the criteria variables of singular behavioural acts in order to obtain a more general classifications of behaviour (multiple act criteria, see Fischbein and Ajzen 1974). The two hierarchically organized structures of personality constructs and behavioural constructs can then be combined symmetrically as is shown in figure 1. The lens-like structure in which many variables on different levels of aggregation refer symmetrically to each other is called Brunswick´s lens model (Brunswick 1956). Based on a set of psychophysiological data Wittman and Schmidt (1983) could demonstrate the usefulness of this model. It indeed enables one to obtain from a given set of personality data more reliable predictions of human behaviour.

In the parapsychological context Mischo and Wittmann (1980, 1981, Mischo 1974) applied this method for the post hoc evaluation of a multivariate ESP experiment. It was shown that the application of stepwise multiple regression analysis which also accounted for non linear dependencies could enhance dramatically the portion of explained variance. Timm (1982,1983), however, observed that the use of this special method may be questionable because statistical selection errors could occur.

Indeed, it is true that the multivariate approach loses much of its power because of technical problems of the statistical tests applied. The application of available statistical tests is generally restricted by certain parametric assumptions which may not be justified in practice. Moreover, in parapsychology no theoretical framework is available from which a strategy could be derived how to aggregate and to select the relevant variables (see 1. and 2. in chapter 1). It is obvious that such a strategy is essential for the multivariate method, since without it any real correlation between psychological and physical variables (which may indicate a PK effect) could be diluted by a mass of irrelevant variables which were measured simply because one did not know which variables were relevant and which were not. This simple example indicates that multivariate methods such as factor analysis, cluster analysis and others should not be considered as a mechanical method to gain scientific knowledge! The multivariate method might be a good method to describe very complex systems but

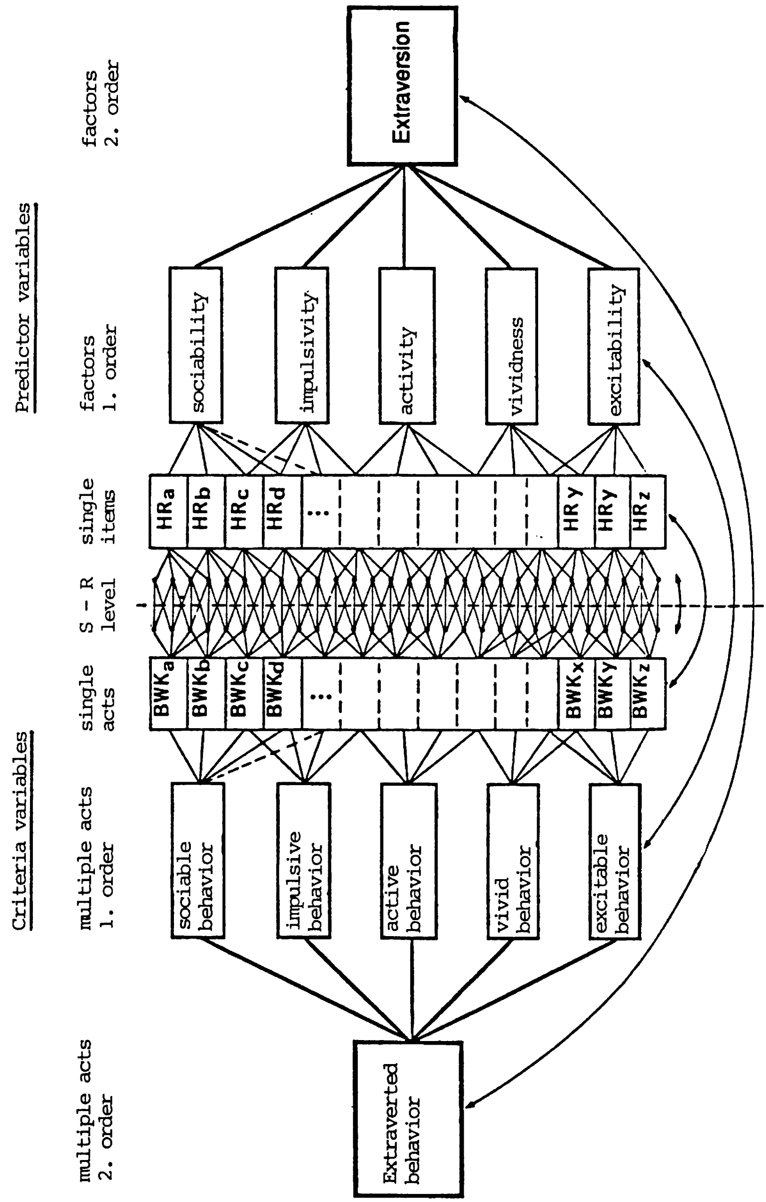

FIGURE 1

Brunswick's hierarchical lense model to illustrate the symmetry of predictor and criteria variables in Eysenck's theory of extraversion only in the context of good theoretical models (see also Fahrenberg 1982).

4. THE EXPERIMENTAL DESIGN

The study consists of two parts which are quite different from a methodological point of view. The less rigorous one is called the pilot study, the other, the main study. The pilot study, however, was not performed prior to the main study, but parallel to it. In the main study the entire procedure was standardized. The subjects were unselected persons from Freiburg and vicinity. The experiment was announced at the University and in a local newspaper. It was emphasized that also persons who were sceptical about their psychic abilities or about paranormal claims were welcome to participate in the experiment. The subjects for the pilot study were selected persons, who either claimed to possess psychic abilities or who, for instance, were involved in spontaneous paranormal manifestations such as RSPK cases.

The experimental conditions of the pilot study could be changed, and this was often done in accordance with the wishes of the subjects. The number of runs was not predetermined and the psychological and situation variables were described only qualitatively on the basis of the experimenters or other persons judgement. The purpose of the pilot study was not to prove the existence of PK but to optimize the experimental conditions in such a way that the PK effect obtained would be as large as possible - at least to the experimenter´s own judgement. For the evaluation of the pilot study only those runs were selected which exhibited - to the judgment of the experimenter - successful PK manifestation. It should be noted, however, that some of these pilot experiments fulfilled the usual methodological requirements of standard PK experiments. In part III some examples will be described in more detail. The more or less subjective method of the pilot study was needed in order to obtain some heuristics about the structure of the PK process. It was especially aimed at finding a specific variable, which could serve as a tracer for PK. Such a tracer variable could then be used for the evaluation of the main experiment as a special (physical) criterion variable. It was assumed that such a tracer variable would exhibit higher correlations to one or more relevant psychological variables than all the other variables. For this purpose, however, it was necessary that at least the physical structure of the PK apparatus was the same in both the pilot and the main study. To gain possible heuristics for future experiments occasionally certain parameters of the target production process were changed. Moreover, in the pilot study occasionally alternative types of display were used.

The main experiment was carried out in a nice old villa. The room for the subjects was relatively small but had comfortable furniture and created a friendly atmosphere. The computer and the random generator were placed in an adjacent room as shown in figure 2. Only the display was put on a table in the subject´s room. In some sessions of the pilot study the subject and display were located inside a shielded room with an attenuation of 60 dB in the range of 10 kHz to 30 GHz. The display data from the PK apparatus were fed into the shielded room by an optical channel.

4.1 The PK apparatus

The apparatus consists of three components:

1.) A computer system which controls the
2.) random number generator (RNG) which presents its results on a
3.) display.

The technical details of the computer system (hardware) and the software system are described elsewhere (see Lucadou 1986). It should be mentioned, however, that special precautions were taken to prevent electrical artifacts.

The display was a small black metal box with five buttons and a string of 16 red lights (LEDs) (see figure 3). When the run started the lower seven lights were switched on. If the next trial was a ´1´ the next light at the top of the row was lit, if a ´0´ appeared the light on the top was switched off. Thus according to the binary random sequence produced by the RNG the light fluctuated up and down similar to a thermometer column. If the column of the lights reached the top of the string a further hit could not be detected by the subject. The light column remained at the top until the next ´0´ would arrive. Respectively, if all lights (except the lowest one which served as a control to indicate that the apparatus was in operation) were off (the column had reached the bottom of the row) then a further miss (´0´)

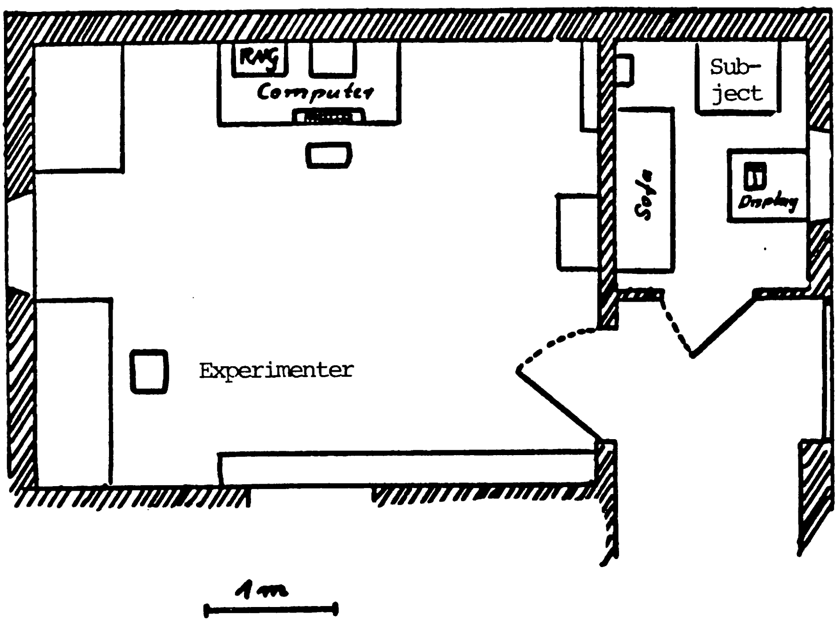

FIGURE 2

Site of the laboratory

could not be detected. Thus the subject received only a restricted range of feedback as regards the deviation from the mean value. The rate of target generation and (real time) display presentation was about 10 per second. During the pilot study an additional acoustical feedback could be presented in which the pitch of a tone varied according to the random sequence. Moreover, the system also included an optical display on a video screen, where ´lights´ were arranged similar to the display used by H. Schmidt (see Schmidt 1970). At some occasions this display was used in the pilot study.

The random generator is shown schematically in figure 4. The random source is a small cylindric radio-active probe (strontium 90) with an activity of 0.045 mCi contained in a plastic block of 40x35x50 mm in size. In each of five holes around the radio-active source 5 Geiger-Mueller-tubes (ZP1310) are placed. The plastic block with the radio-active source and the five Geiger-Mueller-tubes are electrically shielded by a container of aluminium. The pulses coming from each of the five counters are fed into five independent electronic circuits (channels) each generating two independent binary random sequences by two different methods. They are called Schmidtian and Markowian random generators. The Schmidtian RNG produces real random sequences and the Markowian produces Markow-sequences. Thus, the random generator produces per time interval 10 independent parallel random events, 5 Schmidt and 5 Markow events. In figure 4 only one channel is shown. Only one of these 10 parallel random sequences is presented on the display, but all are recorded in a computer file. The independence of the five channels is guaranteed by the fact that a particle which is emitted from the source can only be detected by one of the five Geiger-Mueller-tubes.

The simple question behind this arrangement is: What will the other four detectors show if by PK a deviation from chance is produced on the channel which is connected to the display? Before discussing this question in more detail it is necessary to describe the process of target-generation.

4.2 Methods of target generation

The Schmidt RNG is based on the same principle as was applied by H. Schmidt (see Schmidt 1970) in his experiments. An electronic switch (flip-flop) is triggered by a ´high´ frequency (5 MHz) clock (silicium

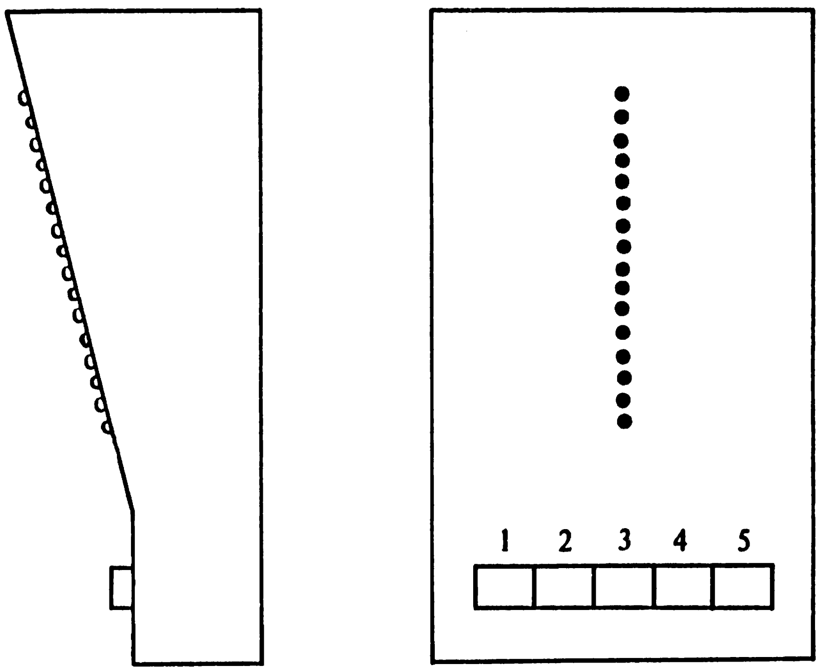

FIGURE 3 The display

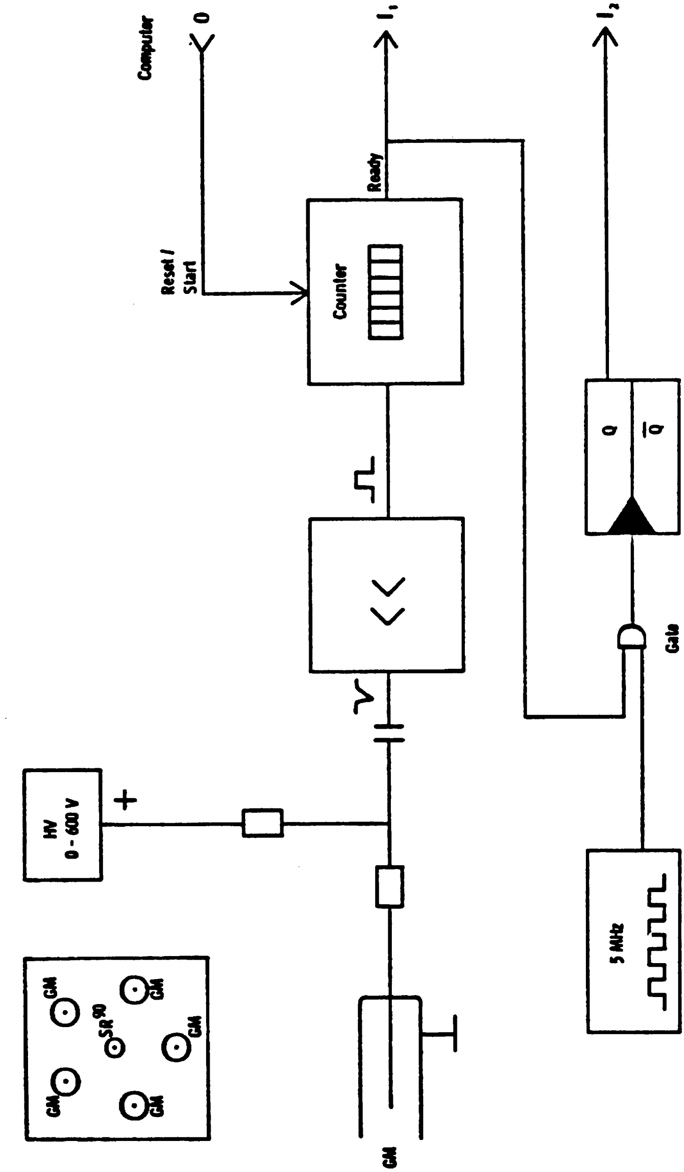

FIGURE 4 Block diagram of the random generator

oscillator). This frequency can be considered high in relation to the frequency of target generation. Whenever a pulse from the Geiger-Mueller detector arrives this process is stopped. The state in which the flip-flop remains defines a hit ´1´ or a miss ´0´ of the sequence. Since the length of the time intervals between two pulses of the radio-active source is random (stochastically independent from the previous one) one can prove that the result will be an equally distributed random sequence (see Schmidt 1977). The time interval between the pulses defines the target generation rate. In our experiment the trial rate could be adjusted in two different ways: 1. Since the computer started each target generation procedure by a start signal the generation rate could be controlled by the computer. The start signal was the same for all five channels. 2. The pulses from the Geiger-Mueller-tubes (2,5-3 kHz) were not directly submitted to the gate of the flip-flops, but they were first sampled in a counter. After having reached a given number of counts a ready signal was produced to stop the flip-flop. The number of counts could be set for each channel independently. Only if a ready signal had arrived at all five channels the next start signal from the computer could be given. Thus the target generation rate also depends on the setting of the counters.

It is obvious that it is not possible with the Schmidt´s RNGs to localize a possible PK effect in relation to the radio-active source, because there is no correlation between a hit (or a miss) and the decay rate of the source. If one wants to know whether the decay rate of the source can be altered by PK it is necessary to use a different technique of target generation.

This new method also uses the ready-signal from the counter which stops the flip-flop, but in this case also the time interval between the start and the ready-signal is measured by an internal clock of the computer. This time interval is a direct (inverse) measure of the decay rate of the radio-active source. Here we call it ´counting time´. Table 1 gives the standard setting of the five channels during the main study together with the typical average counting times of each channel. Different settings of the counters were chosen in order to find out whether an accumulation of the PK-effect during the counting interval could be found. With this setting a display rate of about 10 Hz was achieved. The counting times are measured in units of internal computer clock cycles. After the arrival of the last ready signal of the five channels the computer generates a new start signal to all channels simultaneously. The counting time (in units of computer cycles) of the last (slowest) counter is also stored on the computerfile for each target generation together with the set of the 10 parallel events (as discussed above).

TABLE 1
Setting of the counters of the channels
and average counting times in computer clock cycles

Channel	setting of the counter	counting time
1	152	1380
2	192	1620
3	112	1110
4	32	340
5	72	690

The binary Markow random sequence is generated for each channel in the following way:

The counting time Zi(t) at a certain instance t is compared with the countingtime Zi(t-1) of the previous trial.

If Zi(t) < Zi(t-1) then a miss ´0´ is generated
If Zi(t) > Zi(t-1) then a hit ´1´ is generated
If Zi(t) = Zi(t-1) then the target generation will be repeated, which means that this case will be ignored.

Since the variance of Zi: si = 9*SQRT(MEAN(Zi)), is large the last case occurs rarely.

It can be shown (see Lucadou 1986) that the resulting random sequence is a Markow-chain of the first order which is specified by the following transition matrix:

$$(pij) = \begin{pmatrix} p00 & p01 \\ p10 & p11 \end{pmatrix} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}$$

p00 describes the probability to obtain a ´0´ after a previous ´0´, and so on. It is remarkable that these values hold exactly and do not depend on the form of the distribution of the initial random process (which in our case is a Poisson distribution). The only requirement which must be fulfilled is that the single events of the source are stochastically independent. The transition matrix completely describes the Markow-chain. Especially, it follows that the (ergodic) probabilities of a hit and a miss in the whole sequence are equal:

$p0 = 1/2; \quad p1 = 1/2$

The distribution p(n,T) of hits T in such a Markow-chain of a given length n can be calculated from an algorithm (see Lucadou 1986). For $n > 10$ this distribution is very similar to a normal distribution (Gaussian) with $p = 1/2$ and $s = SQRT(n/12+1/6)$. For $n = 600$ which was the standardized run length for the whole study one can even use $s = SQRT(n/12)$. Thus a CR-value for the Markow-chains can simply be calculated in an equivalent way as with the normal random sequences of the Schmidt´s RNGs:

$CR = (T - n/2)/SQRT(n/12)$
T: numbers of hits, n: length of sequence.

4.3 The correlation technique

It is obvious, that with the Markow-sequences a one to one correlation between the binary events (hits and misses) and the decay rate of the source is given. Assuming that the decay rate of the source declines at least over a period of .2 seconds (two trials) this would produce two hits. Respectively, if a series of hits has occurred one can conclude that the decay rate must have declined during this period of target generation. If it is true that the decay rate has declined, then of course the other four counters must also have produced a series of hits. Thus, it simply needs a comparison with the four remaining Markow-sequences to settle the question whether PK can influence the decay rate of a radio-active source. If this hypothesis has to be refuted, however, one could still imagine other processes, such as an influence on the angular distribution of the radio-active emission of the source. In this case one would expect, for instance, that only two of the Markow-chains look similar. If this also would show not to be the case then one can conclude that no common influence has acted on the different counters.

In order to compare two or more Markow-sequences a rather sensitive technique called correlation technique can be used.

A cross-correlation function of two numerical series i and j of length n is given by:

$$\phi ij(k) = 1/n \sum_{1}^{n} gi(1)*gj(1+k)$$

gi: element series i, either 0 or 1
gj: element series j, either 0 or 1

This expression is a measure for the similarity of the two sequences when they are shifted against each other over k elements. The distribution of ϕij values for different k is an indicator for the similarity between the two series. For the auto-correlation function, only one sequence is used, hence i and j are identical. Thus, the auto-correlation function is a measure for the similarity of a sequence with itself.

If the Markow-sequences of the four non-display channels would become more similar due to a common PK influence on the source, one would expect the cross-correlation function between each non-display sequence and the display also to become more similar to the auto-correlation function of the display sequence. If these sequences become identical with the display sequence the cross-correlation function will be equal to the auto-correlation function of the display sequence defined as:

$$\phi i(k) = 1/n \sum_{1}^{n} gi(1)*gi(1+k)$$

gi(1): element 1 of display sequence.
gi(1+k): element at distance k of element gi(1) of display sequence
i is the number of the display channel.

The auto-correlation function of the Markow-chain is known under the null hypothesis. Therefore the similarity or difference between the auto-correlation function of the display sequence and the cross-correlation function of any other sequence with the latter yields a good measure for a possible common PK influence on the five channels. (Because of the nature of the Schmidtian random process it is not useful to study cross-correlation functions of these sequences).

In figure 5a,b examples of the auto-correlation function and the four cross-correlation functions are shown. In figure 5a there is no common influence (null hypothesis) and thus there is no similarity between these functions. In figure 5b, however, one can see that the functions look very similar although they are derived from random processes which do not show a clear similarity if they are compared directly. This shows that the auto- and cross-correlation method can reveal similarities which cannot be detected by comparing the original data sequences. One can say that in the original sequence the common influence (or signal) is buried in the noise of the source.

To simulate a common influence the source was mechanically moved in an irregular way along the symmetry axis. This causes a common varying pulse rate in the five Geiger-Mueller-tubes. As a result, the functions became very similar. Figure 5 is based on the data of this simulation study. This demonstrates that the method does not require that the nature of the influence must be known beforehand, fortunately, because in the case of PK that would be very difficult. It should be mentioned, however, that further research is needed to provide a quantitative measure for the accuracy of this method.

If one compares figure 5a with figure 5b it is apparent that the auto-correlation function of the display sequence has changed its form due to the (simulated) influence. Thus a comparison of the actual auto-correlation function with the theoretical auto-correlation function may serve as a sensitive measure for any deviation from pure random behaviour and may be used as a possible tracer.

4.4 The experimental procedure of the main experiment

The experimental procedure is schematically described in figure 6. The subject first had to fill out several questionnaires at home

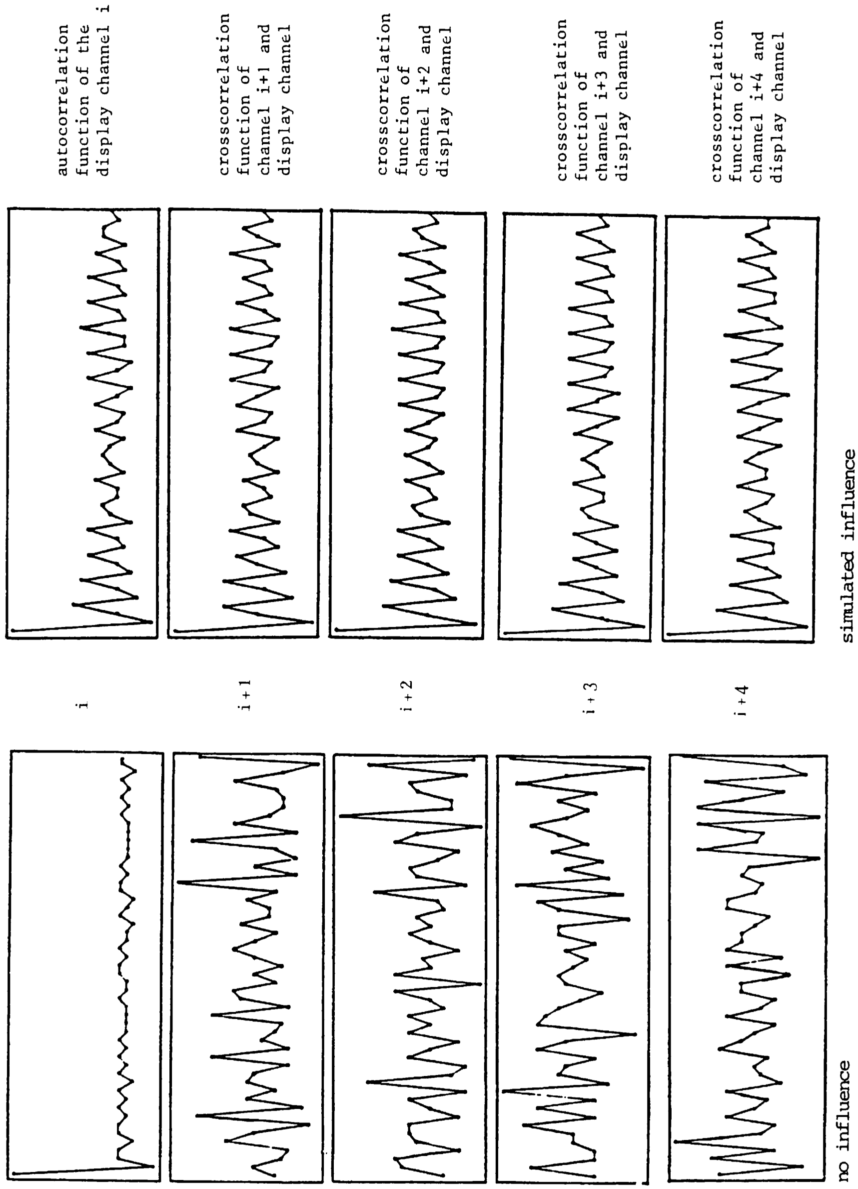

FIGURE 5 Auto- and crosscorrelation function of the five channels with and without simulated influence

before he or she was allowed to take part in the experimental session. These questionnaires were collected before the session for that subject was held. The questionnaires measure trait variables: 1. sheep-goat. 2. personality traits which have physiological relevance (FPI, Fahrenberg, Selg, Hampel 1978). 3. locus of control (IPC, according to Rotter)(Krampen 1981). Immediately before the experimental session started the subject had to fill out another questionnaire (EWL) on momentary state of mood (Jahnke and Debus 1978). Then the subject performed four feedback runs (runs in which the data of the display channel are shown) each lasting about one minute. After each run there was a pause of one minute. During the pause a non-feedback run was carried out in which all conditions were the same except that without the subject being aware of it the display was switched off. Each feedback run was started by the subject by pushing one of five buttons. The button chosen gave the subject's estimation about his or her momentary 'PK abilities' (confidence score). The non-feedback run was also started by the subject pushing one of the buttons. In this case, however, the subject was asked by choosing the appropriate button to express an assessment of the success which was accomplished in the feedback-run just finished. During the pause the subject received true feedback on his or her previous results by a number of lights which were switched on in relationship to the positive CR value of the deviation from chance.

The experimental conditions within the runs were changed on a double blind basis. Each feedback run was followed by a non-feedback run under exactly the same conditions except for the feedback. By means of a built-in (pseudo-) random generator it was decided whether the first run started with a Markowian or a Schmidtian sequence at the display and which of channel 1-5 was coupled to the display (to avoid a systematic bias of a special channel). All further experimental conditions were fixed by this choice in the following way: Each Schmidtian feedback/non-feedback run pair was followed by a Markowian pair and vice versa. The pseudo-random RND(x) function decided the entry point in the following cyclic sequence:

1M,1S,2M,2S,3M,3S,4M,4S,5M,5S

The number indicates the channel and M respectively S the type of the RNG. For instance, if the RNG selects 3M, this means that channel 3 provides the data for the display, and that the first feedback and non-feedback run is of the Markow type.

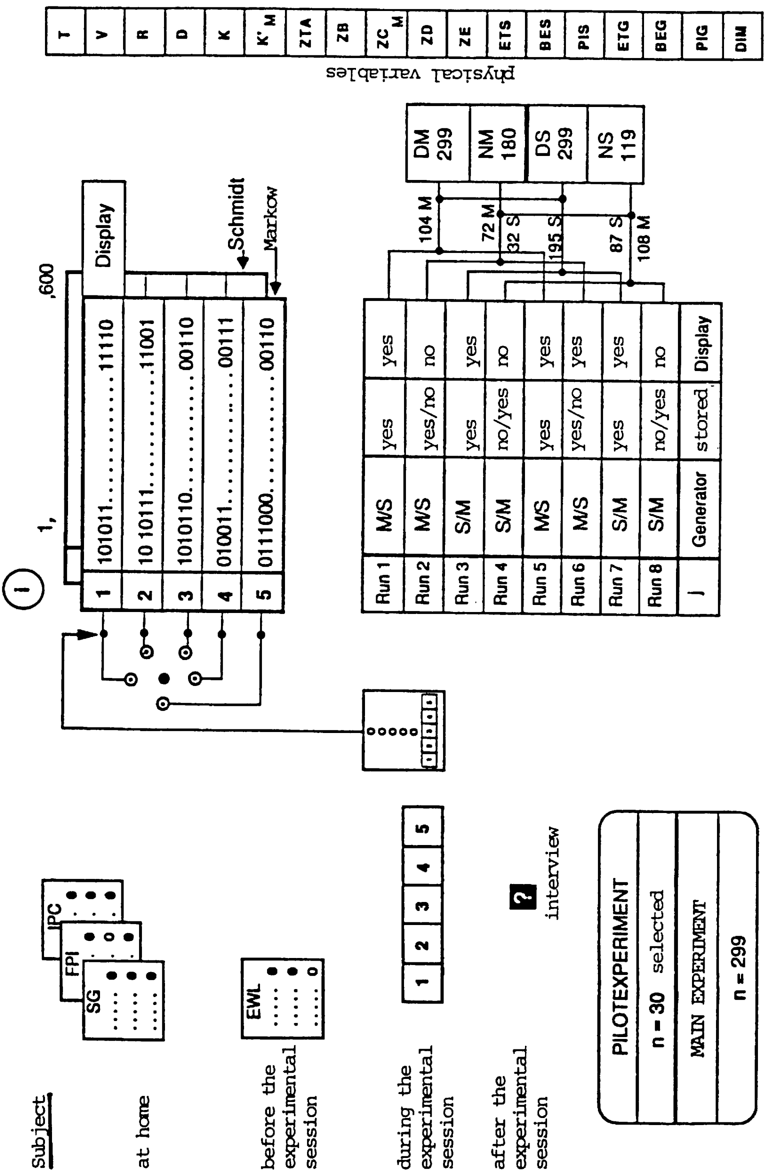

FIGURE 6 The experimental design

For practical reasons (storage capacity) only half of the non-feedback runs were stored in computer files and were included in the final evaluation. In any case, however, the total number of hits of each run was documented on a paper print out which later also served as a control for data security. The decision which non-feedback runs were stored was also made on a double blind basis by the RND(x) function.

Unfortunately, at the end of the experiment it turned out that the group frequencies for the different run conditions were rather unequal both as regards selection of channels for the display as well as for the selection of data storage of the Markowian and the Schmidtian non-feedback series. This was not planned and not known to the experimenter beforehand. It turned out that the pseudo-random function RND(x) always started with the same seed conditions if the program was started. This caused a rather strong bias of the RND(x) function. This disadvantage made it impossible to evaluate the experiment on the basis of single runs, since some groups were too small. Therefore, it was decided to neglect the order and the channel of the single runs and only to use four groups under the following conditions:

1. Display runs of the Markowian RNG (DM), 299 subjects
2. Non-display runs of the Markowian RNG (NM), 180 subjects (E = 149.5)
3. Display runs of the Schmidtian RNG (DS), 299 subjects
4. Non-display runs of the Schmidtian RNG (NS) 119 Subjects (E = 149.5)

Since each subject did 8 runs under these four conditions, each subject contributed two runs to each condition DM, NM, DS and NS. The physical variables (described in the next paragraph), however, can only be defined on the basis of single runs. Therefore, it was necessary to add the corresponding variables for the two runs for each subject.

In total, 299 subjects from which a complete set of psychological data could be obtained were included in the evaluation. The criterion for completeness was that not more than 5% missing data in the questionnaires were allowed. These circumstances resulted in the group frequencies shown above.

4.5 The physical variables

For the physical side of Brunswick's lens a set of statistical quantities and/or variables of the random sequences were used. Part of these variables are statistical quantities which are traditionally used in parapsychological research and of which the distributions under the null hypothesis are known. Other variables, however, were newly defined in order to test specific hypothetical assumptions. Some of them exhibit certain characteristics which were suggested by the pilot study, of others, the characteristics were derived from theoretical models. Most of them can be defined for both the Schmidtian and the Markowian sequences, while two can only be defined for the Markow-chains. It turned out that the physical variables of the Schmidtian and the Markowian sequences could not be pooled together, partly, because they had a different statistical distribution and partly, because they even have a somewhat different meaning. In the following these variables are described:

1. The number of hits:

$$T = \sum_{1}^{600} x \qquad x = (0,1)$$

This is simply the number of '1' in the random sequence produced by the RNG (run-lenghth = 600). In order to compare T of the Schmidtian RNG (Ts) with T of the Markowian (Tm) one, Ts can be transformed by the following formula:

T = ((Ts-300)/SQRT(3))+300

After this transformation the CR value can be calculated for both sequences with the same formula:

CR = (T-TO)/SQRT(n/12), TO is the expectation value of T

2. The hit variance:

$$V = 1/(m-1) \sum_{1=1}^{m} \left(\sum^{20} x - \mathrm{mean}(x) \right)^{**}2$$

measures the fluctuation of hits based on the run divided in 30 parts of equal length. Its expectation value under the null hypothesis is 1.833 for the Markowian RNG and 5.000 for the Schmidtian one.

3. The run score variance

$$R = \sum_{l=1}^{n} CR_{20}^{2}$$

is just a different measure for fluctuations based on a run split up in 30 parts. It is chi-square distributed with 29 degrees of freedom for both Markowian and Schmidtian runs.

4. The Kolmogorow-Smirnov-test quantity

$$D = SQRT(m) * MAX(ABS(F0(x) - F(x)))_{20}$$

gives a measure for any deviation from the hit-distribution of the 30 partial runs under the null hypothesis. F(x) is the cumulative hit-distribution, F0(x) is the cumulative distribution under the null hypothesis (n=30 and p=1/2). The distribution of this test quantity is known (Claus and Ebner 1979). It should detect even such deviations from the null hypothesis which leave both mean value and variance unchanged. In practice, however, it turns out that this statistic is not very sensitive.

5. The auto-correlation test quantity

$$K = \sum_{k=1}^{100} (\phi_{ii}(k) - \phi(k))**2/\phi(k)$$

measures the deviation of the actual auto-correlation function $\phi ii(k)$ from the theoretical auto-correlation function $\phi(k)$ (under the null hypothesis) as discussed in paragraph 4.3. Only the first hundred steps of the auto-correlation function were used. K is a measure for stochastical dependencies during the PK process (non-randomness). Its theoretical distribution is not known.

6. The cross-correlation test quantity

$$K' = \sum_{k=1}^{100} (\phi_{ij}(k) - \phi_{ii}(k))**2/\phi_{ii}(k)$$

This value is only defined for the Markowian sequences. It is a measure for the similarity of the 'parallel' Markowian channels (i,j) to the one connected to the display (i). Its theoretical distribution is not known.

7. Deviations in the target generation time

$$ZTA = SQRT(m)*MAX(ABS(F(Zmax)-F(Z)max_{Gauss})_{20})$$

F(Zmax) is the cumulative distribution of the generation rate.

This variable measures fluctuations in the target generation rate. The target generation time of each set of the 10 parallel events is the time which is needed to fill all five counters. As discussed in paragraph 4.2 the target generation time depends directly on the decay rate of the radio-active source (or the pulse rate of the Geiger-Mueller-tube). Any change in the decay rate would of course also cause a change in the target generation time. It turned out, however, that the pulse rate of the Geiger-Mueller tubes was not sufficiently stable over the whole period of 1.5 years (due to a change in the sensitivity of the tubes) to use the mean-value of target generation time as a measure for deviations. Thus the Kolmogorov-Smirnow test was used to measure deviations from the theoretical normal distribution (F(Z)max,Gauss) of the generation time with the mean-value and variance of the actual measured run. The theoretical distribution is obtained by taking the observed mean-value and variance and to construct a normal distribution fitting these parameters. The variable ZTA therefore mainly measures short fluctuations of the generation-time during a run. Smooth shifts cannot be detected. The distribution of this test quantity is the same as for test quantity D (4).

8. A measure for deviations in the number of hits that does not cause deviations in the variance and vice versa.

$$ZB = ((T-TO)+(V-VO))/(T-TO)*(V-VO)$$

VO is expected variance.

This variable only accounts for such deviations in the hit-rate (T-TO) which do not cause a deviation in the variance (V-VO) of the hit-distribution and vice versa. This measure was defined because during the pilot study the experimenter got the impression that some subjects tended to obtain deviations from the mean-value of hits but no deviation in the hit-variance while others tended to obtain only deviations in the variance. From a theoretical point of view this seems to be an interesting question, since for normal distributions mean-value and variance are independent. The theoretical distribution of ZB is not known.

9. Fluctuations in the display channel

$$ZC = (T-TO) * \sum_{i}^{4} K'$$

is also a variable which is defined based on certain heuristical findings of the pilot study. SUM(K′) is the sum of all four cross-correlation test values (see 6.) of the four parallel channels which are not displayed. If this sum declines this indicates a common influence on all channels since the differences (K′ values) between the cross-correlation functions of the non-display channels and the auto-correlation function of the display channel become smaller. If the scattergrams of the variables T and SUM(K′) for all (selected) runs of the pilot study under feedback and non-feedback condition are compared with each other one gets the impression that under feedback conditions runs with high deviations from TO also show high values in SUM(K′). ZC is a measure which puts more weight on those hits which do not coincide with a joint fluctuation of the source (small values in SUM(K′)) than on single fluctuations in the display channel. ZC is only definable for Markowian sequences, because cross-correlation functions are only meaningful for the Markowian RNGs. The theoretical distribution of ZC is not known.

10. Another variable to distinguish such different sources of fluctuations is:

$$ZD = (T-TO)/ZTA$$

since ZTA is also a measure of a joint fluctuation (see 7.). It is however much less accurate since fluctuations in a single channel can also contribute to ZTA. It was also calculated for the Schmidtian sequences. In this case, however its meaning is not clear because one does not know whether and how the generation time has an influence on the target generation of the Schmidtian sequences. The theoretical distribution of ZD is not known.

11. Chaotic influence

$$ZE = \sum_{i}^{5} Ki$$

ZE is the sum of all five auto-correlation test quantities Ki (see 5.) of the five parallel channels active in the run. It was observed in the pilot study that in the case of one subject it seemed as if all five channels were influenced in a somewhat chaotic way so that the variable ZE yielded high values. The interpretation of this variable is not very clear in relation to a possible underlying model. It simply states that a certain accumulation of irregularities occurred in all five channels. The theoretical distribution is not know.

The remaining 7 variables are different from the previous ones in so far as they were deduced from a theoretical model and without prior investigation in the pilot study. Furthermore they cannot be defined by a simple formula like the previous ones but they are operators (recursive algorithms) which act on the whole random sequence. In contrast to the previous variables they not only describe physical features of the random process, but also reflect certain ´psychological´ assumptions about the effect of what the subject actually perceives on the display. The theoretical background will be given in paragraph 7. (Details of the algorithms are described in Lucadou 1986). Here only the ´meaning´ of the variables is shortly described. These variables can be defined for both types of RNGs. Their theoretical distributions are not known.

12. ETS

This variable is a measure for the amplitude of fluctuations of the column of light which is presented on the display. Only ´positive´ fluctuations (in the correct direction according to the instruction) are counted. A ´positive´ fluctuation is defined by a sequence of hits occurring in an partial run of a dynamic length which changes for each trial. This length is given by the number of steps from the present position of the light column to the top of the row. If the number of hits in this length exceeds a certain threshold (CR=2) the sequence is regarded as a ´positive´ fluctuation. ETS is the sum of the positive fluctuations of all partial runs.

13. BES

This variable is a measure for the duration of a display fluctuation described in 12.

14. PIS = / ETS*BES

is the floating product of ETS and BES.

15. ETG

is the same variable as ETS, however, for fluctuations into the ´wrong´ direction.

16. BEG

is the same variable as BES, however, for the duration of such ´wrong´ fluctuations.

17. PIG = / ETG*BEG

is the floating product of ETG and BEG.

18. DIM

is a variable which takes into account the display-function as it is described in paragraph 4.1. It measures the ´visual´ success of the subject according to his or her instruction. DIM is the sum of the ´free intervals´ of all trials of the run. The ´free interval´ is defined as the distance between the top of the column of light and the top of the row. If the subject is able often to push or to keep the column to the top of the row during the run DIM will be a small value. If not, DIM will be large. Thus, it is inversely related to the success of the subject. One could also say that this variable describes the ´meaning´ of the display to the subject. The meaning can only be defined in the context of the instruction as will be discussed in paragraph 7.

5. A METHOD TO FIND SIGNALS IN RANDOM SEQUENCES

Especially the physical variables nr. 1 to nr. 11 were designed to serve as a tracer of a possible PK influence. However, from the evaluation of the pilot study it appeared that they are not useful for this purpose. No tracer was found in the pilot study. These variables are defined on the basis of specific empirical or theoretical models. Therefore, they might not be sufficiently general to serve as a method to detect arbitrary or unknown signals in a random sequence. A signal embedded in a random sequence means that there is an external or internal, but independent, process that modifies or modulates the pure random process in such a way that it is no longer purely random. This implies that it must be possible to distinguish at least with a certain probability a pure chance event from a non chance event of the process. This should be possible even for those sequences which do not show significant deviations in a given statistical test value. Practically, this means that a certain filter has to be applied to filter out the initial signal from the ´background noise´ of the random sequence. As an example, a sinusoidal signal may serve which is superimposed on a random sequence in such a way that the mean value of the distribution remains unchanged. Signal enhancement techniques (Ehrenstrasser 1974) have been developed to filter out such signals from the background. The main problem is that the efficiency of such methods depend very much on additional knowledge about the hidden signal such as frequency, phase, or amplitude. In parapsychology, however, one does not know anything about the form of the PK signal which could be used to filter out the signal. Therefore, one would need a ´general´ filter to filter out the PK influence.

This constitutes a difficult theoretical problem, since a formulation of a ´general´ filter cannot be given. This is due to the fact, that we cannot provide a logically complete description of nature (see Lucadou and Kornwachs 1974). Of course, one has to rule out that a possible signal is purely random itself. If not, it could no more be regarded as a signal because there would not be any criterion left to distinguish it from the original random process.

From a pragmatic point of view it would be sufficient to find a filter which is ´general enough´ to detect a ´PK signal´ in those sequences which can be regarded as successful PK manifestations according to the criteria usually applied in parapsychological research.

The rather general method I finally found was based on the idea that only combinations of different statistical test values could deliver sufficient information about a possible underlying non-random process. In analogy to the phase space concept in (quantum) physics, it was assumed, that pairs of independent (canonical or orthogonal) variables may be sufficient to describe the state of a given system (see Lucadou 1974). (In quantum physics pairs of complementary observables deliver the maximal information which is available on the state of the system). Along these lines one could argue that certain PK signals may be classified by their simultaneous deviation in hit-rate and variance. If PK could be described by these two independent variables one would expect that a scattergram showing the correlation between hit rate and variance would exhibit certain regions which might mainly be reserved for ´PK runs´ (see also variable ZB). However, from scattergrams based on successful PK runs of the pilot study and on data of simulation studies it followed that the pair hit-rate (T) and variance (V) was not sufficiently sensitive to detect even strong signals.

By trial and error I finally found a pair of variables that turned out to ´react´ rather sensitively to both regular and irregular simulated signals. These two variables are the hit-rate (T) and auto-correlation test value (K). In figure 7 the correlation between these two variables is presented. The data are based on control runs. The larger dark points in the scattergram indicate those runs in which a non-random sequence was superimposed, generated by moving the radio-active source along the symmetry axis (as described in paragraph 4.3). The other points are based on true RNG runs. For convenience the quadratic correlation can be linearized. If a point in this scattergram lies outside the distribution of the curve one can conclude that it contains a signal, even in the case that one or even both of the two test values are inside the ´allowed´ region of the null hypothesis.

It should be mentioned, however, that as yet the sensitivity of this method is not quantified. This would imply a categorization of signals and their effect on the method. This must be the subject of further research. From figure 7 it seems plausible, however, that the method is especially sensitive in the case of runs with large deviations from

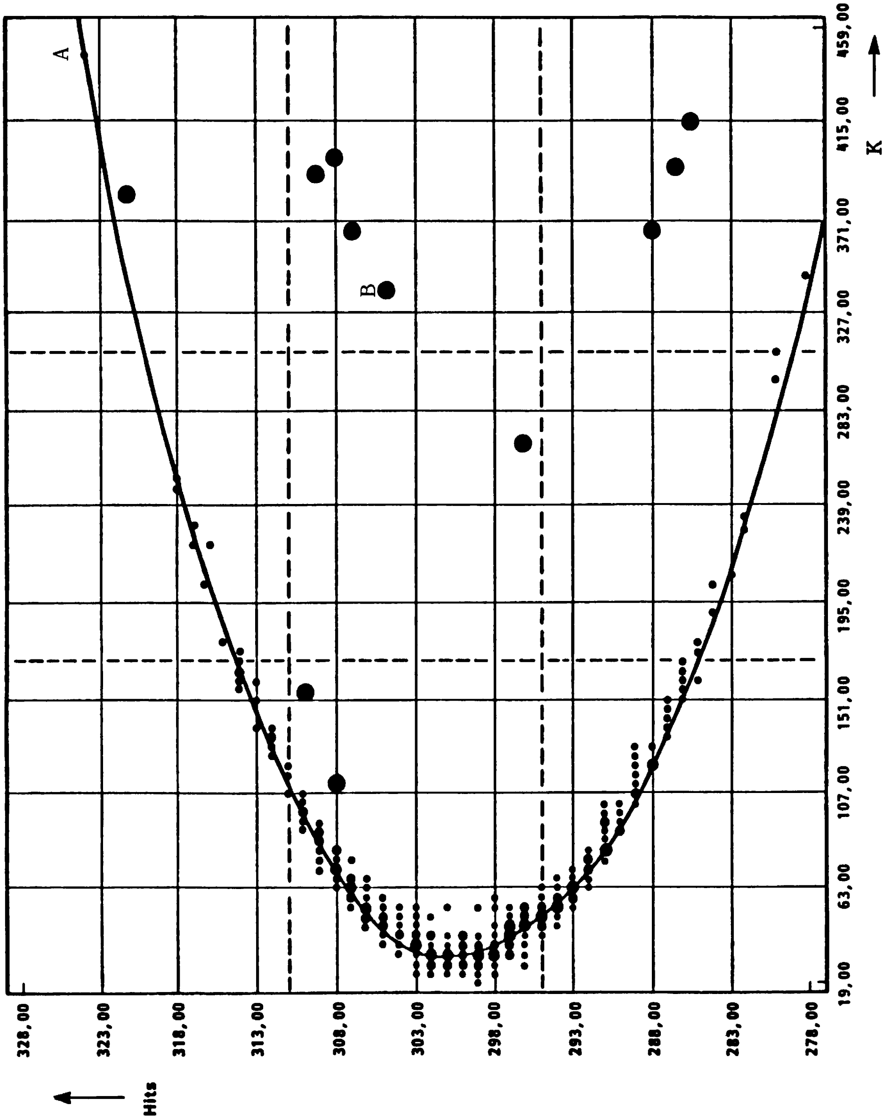

FIGURE 7 Correlation between the hitrate and the autocorrelation test quantity

the expected hit-rate (T) value.

Other combinations of variables were also investigated, but they did not yield such a clear distinction between random runs and those with a simulated influence.

The method described can be considered as a kind of state-space representation in which a pure random process can only occupy certain areas in this space (within a certain probability). It is important to note that this method can also be used to show that a certain run which may exhibit a significant deviation from one or the other expectation values (for instance A in figure 7) must still be considered as a random fluctuation if it lies on the curve of the scattergram. Even seemingly paradoxical situations can occur in which, for instance, a run without significant deviations appears to contain a signal (see B in figure 7) whereas another one with highly significant deviations (see A in figure 7) must nevertheless be considered a random fluctuation without a real signal in it. Therefore, this method is very useful to test the tracer hypothesis even in the case of run selection as it was done in the pilot study. It was applied both in the pilot study as well as in the main experiment.

6. PROVISIONS AGAINST FRAUD AND ARTEFACTS

In parapsychology one has to consider the possibility of manipulation by subjects which could lead to artificial results. In this study two different sources of manipulations must be considered: 1. Manipulations by the subject on the psychological data and 2. Manipulations by the subject on the equipment. The first case is a general problem associated with the use of psychological questionnaires. One can never be certain that subjects fill out a questionnaire honestly. Very often persons tend to present a too favorable or idealized image of themselves in such questionnaires. In some questionnaires, however, there are some internal criteria which could indicate whether the questionnaire is filled out honestly or not. The FPI applied in our study includes a special scale called ´openness´ which measures to some extent the ´honesty´ of the subject. In any case a ´manipulation´ by a subject on psychological data can not contribute to an unwarranted rejection of the null hypothesis in favour of psi, since it was taken care that the questionnaires were returned to the experimenter before the experimental session was carried out. As was shown by Timm (1979), even in the case that the distribution of the psychological data would be highly biased under the null hypothesis, a correlation with the independent random sequences which are produced during the experimental session cannot occur. If the null hypothesis can be rejected, any manipulation of the psychological data by the subject can only increase the error variance and will lead to a decrease of a supposed underlying correlation. Thus, any manipulation on the psychological data by the subjects would be conservative to an assumed PK correlation.

More serious considerations deserve the possibility of manipulations of equipment during the sessions. Hansel (1980), for instance, argued that with some of Schmidt´s experiments such manipulation could have been possible. In the present study, however, with mainly unselected subjects, in which only correlations between psychological and random variables are investigated, such an argument loses much of its power. Even if some of the subjects would have tried to manipulate the equipment they could not have done it in a purposeful way since they did not know the underlying experimental hypothesis.

Nevertheless special precautions were taken to prevent any manipulation. The random-generator and the whole computer system were placed in a separate room (the experimenter´s room) and the subject had no access to any part of the system. Moreover, during the experimental sessions the keyboard of the computer was locked and the subject was not allowed to enter the experimenter´s room without the experimenter being present. The only possible manipulation of the subject could have been performed on the display box, for instance, by opening the screws and/or disconnecting some electrical lines (as it was argued by Hansel). Such manipulations, however, could not lead to any malfunctioning of the RNGs since the display and the RNGs were interfaced by independent input/output lines (with open collector transistors). In the worst case these transistors could be damaged by such a manipulation. This, however, would not lead to a malfunction of the other I/O channels and could be detected very easily. During the entire experimental period of both pilot and main study, no indication was ever found that any subject tried such manipulations.

Furthermore, precautions were taken against possible data manipulation by the experimenter(s). During the experiment, identification numbers and the hit-rate of each run were printed out. This printout also served as a control to prevent recording errors on the floppy disk files. This printout was transfered on punch cards by independent co-workers. Moreover, prior to the evaluation period the whole set of experimental data were stored on a computer tape and kept by an independent institute (IAO, Fraunhofer Gesellschaft) in another city (Stuttgart, West Germany) in such a way that the experimenter had no further access to it.

All hypotheses for the main experiment were formulated in advance and discussed with independent researchers not working in the field of parapsychology. For any interested researcher, a computer tape (containing the raw data and the files for the SPSS routines) can be made available in order to control independently the results which are reported in the following articles.

7. OVERVIEW OF THE FUNDAMENTAL EXPERIMENTAL HYPOTHESES

The hypotheses which were tested in this experiment are quite different from each other in view of their relevance and specificity to certain models and to parapsychology in general. Therefore, it is necessary to describe the relations between these hypotheses. In figure 8 a hierarchical schema of all experimental hypotheses is given. The hypotheses will be discussed in more detail in part II and III. Here only a general overview is presented.

The highest hierarchical level is given by the general PK hypothesis (1). It concerns the question whether physical random processes can show (significant) correlations to independent psychological variables (like personality traits measured before the experimental session) and the independent psychological conditions related to the session (like the instruction to the subject). Nothing is assumed about a specific content of the correlation, nor is a physical process specified which could be responsible for such a correlation. It is not even assumed that feedback may play an essential role. To test this hypothesis on this very general level (namely, a non-feedback PK effect) it would be necessary, however, to know at least the a priori distributions of the physical variables under the null hypothesis. Furthermore it would be necessary to know possible correlations under the null hypothesis among the set of the independent, respectively, the dependent variables in order to calculate the expectation value of the number of accidental correlations under the null hypothesis.

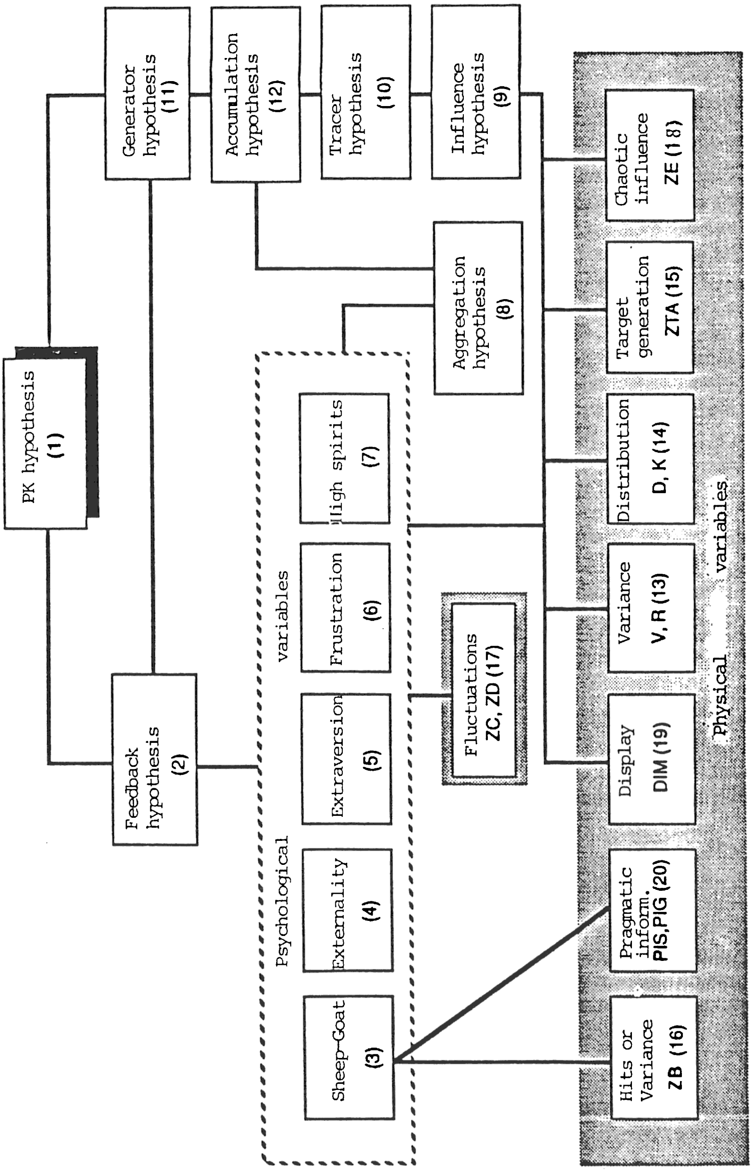

FIGURE 8 Hierarchical schema of the experimental hypotheses

From the description of the variables in paragraph 4.5 it is obvious that this is not possible. Nevertheless, it can be argued that at least some 'strong' correlations between some psychological and physical variables could be indicative of a non-feedback PK effect.

The acceptance of the alternative (PK) hypothesis at a lower hierarchical level of the schema, would obviously presuppose the acceptance of the PK hypothesis at a higher level. Thus, if the feedback hypothesis (2) could be verified at a lower level this would automatically verify the general PK hypothesis.

It may be a matter of discussion whether or not to rank the generator hypothesis (11) at the same level as the feedback hypothesis. One may argue, for instance, that in the observational theories the assumption of the necessity of feedback may be wrong but in this case it could still be possible that PK depends on the nature of the physical process and hence on the type of RNG. On the other hand, even if it should turn out that feedback is essential, PK might still depend on the way of generation of targets. As we have discussed above (paragraph 1), this would allow a decision on the goal oriented model of Schmidt and of Walker's model.

The aggregation hypothesis (8) which corresponds to Brunswick's lens model, results in a question of a general methodological nature. This hypothesis has two aspects associated with the two sides of the lens: It is in question whether the PK correlation can be enhanced 1. by aggregation of the psychological variables or 2. by accumulation of the physical data or even by both of them. These assumptions are, of course, basically inherent to the use of the statistical method, but as I have discussed elsewhere (Lucadou 1984) doubts about it may be reasonable.

The other hypotheses are more specific and can be classified as psychological and physical hypotheses.

The tracer hypothesis (10) supposes that a special 'PK-signal' could be found which allows to discriminate between this signal and the random sequences. It is supposed that such a signal will lead to higher and 'clearer' correlations with the psychological variables, but primarily that it can exhibit a PK effect without the use of psychological variables. The form of such a 'PK-signal' was not specified beforehand. It should be noticed, however, that the methods used here are only suitable to detect certain classes of signals.

Hence it is difficult to refute the hypothesis in a very general sense.

A special case of the tracer hypothesis is the assumption that PK acts on the central radio-active source (or on all five Geiger-Mueller tubes) in such a way that all five counters are affected in a similar way. This assumption is called the 'influence hypothesis' (9).

The psychological hypotheses are derived from previous experiments or from more or less qualitative descriptions of spontaneous cases or from special theoretical models. Thus, for instance, Stanford's model or even qualitative descriptions of RSPK-cases could be used to formulate such psychological hypotheses. The details of these hypotheses are discussed in part II. It is obvious, that the sheep-goat hypothesis and the extraversion hypothesis play a predominant role. On the other hand, for some of the psychological variables measured no specific hypotheses were formulated, because it was the first time they were studied in a PK experiment. It is also questionable whether it is allowed to transfer hypotheses which were formulated in the context of spontaneous cases or ESP-experiments to PK experiments.

A special theoretically derived hypothesis is called 'the hypothesis of pragmatic information' (20). It combines physical and psychological aspects. The model of pragmatic information (which is described in more detail in Lucadou and Kornwachs 1982, Kornwachs & Lucadou 1977, 1979, 1985, 1987) assumes that the meaning of a given information (and not its amount of bits) is the relevant variable in any (para-)psychological experiment. The meaning of a given information, however, cannot be measured by a simple scalar function like the bit rate in Shannon's measure of information. Shannon's measure is only a measure for the (technical) channel capacity. The meaning of the information, however, must also take into account the context, the internal structure and the potential reactions of the system under consideration. A measure of such potential reactions is called 'pragmatic information' (see Weizsaecker 1974). It cannot be operationalized in an absolute way but only in the context of a given system. The first approach for such an operationalization was made in the experiment described here. It should be mentioned, however, that the method used here is not unique and many other (and possibly better) methods may be applied.

It is obvious that the number of hits could serve as a Shannon measure of information of the (potential) PK effect. It is an important weakness that it does not take into account the dynamics and the context of the actual (PK-) process during the experimental session. The instruction to the subject and the form of the display can be regarded as the context of the experiment, and the dynamics is given by the actual random-fluctuations which could be perceived by the subject on the display. The standardized instruction was display oriented and nothing was said about the underlying process. The subject was asked to mentally ´push´ the column of red lights to the top and to keep it there as long as possible. As already discussed in paragraph 4.1 the subject only received relative feedback. From the point of view of pragmatic information the meaning of a single hit in the sequence depends on its actual position on the display (display hypothesis, 19). For instance, if the column had reached the top, a further hit would not have any meaning to the subject. The variable DIM (13.) which weights the hits according to their actual position at the display, thus represents a measure for the meaning of the feedback to the subject in relation to the instruction.

One can argue that certain fluctuations which can be recognized on the display by the subject may have a specific meaning for her or him. Based on this point of view the variables 12. to 17. were constructed. Generally speaking, the model of pragmatic information predicts that such variables which are operationalizations of the pragmatic information will exhibit higher PK correlations to the independent psychological variables than measures which do not contain the aspect of dynamics. It is obvious that the hypothesis of pragmatic information presupposes the feedback hypothesis. On the other hand, a verification of the hypothesis of pragmatic information would strongly support the feedback hypothesis and consequently also the PK hypothesis.

A more precise quantitative formulation of all hypotheses that allow appropriate test statistics will be given in Part II and III. It should be mentioned, however, that this implies a restriction of the more general hypotheses (PK hypothesis, feedback hypothesis and generator hypothesis) formulated above due to the special methods and operationalizations applied in this experiment. Hence, any interpretation of the experimental results in terms of these more general hypotheses includes certain assumptions on the validity of the theoretical models. This clearly illustrates how close theoretical and experimental developments are linked together. Without such links empirical progress would not be possible.

ABSTRACT

The first part of a research report on a multivariate PK experiment is presented. The experiment should shed some light on both physical and psychological aspects of PK. In relation to this problem it is useful to define PK not only as an anomalous high (or low) hit-rate but as a (significant) correlation between psychological and physical variables. Some theoretical implications of this approach are discussed. From the experiences in (personality) psychology it cannot be expected that such correlations are strong. Therefore, it is necessary to combine many independent and dependent variables on the basis of Brunswick's lens model.

Regarding the physical aspects it seem of fundamental importance whether a specific physical variable (tracer) can be found which indicates the occurrence of a PK effect. To this purpose, a special correlation technique was developed which allows to detect signals in random sequences.

The radio-active decay (e-) of a SR-90 source, centered in a circle of five Geiger-Mueller-detectors, served as the random process in the PK experiment. Subjects were 299 unselected persons (mainly students). The display for the subjects consisted of a string of 16 lights (LEDs) performing a column of lights fluctuating up and down, similar to a thermometer column, depending on the outcome of the random process. Two different types of random sequences produced by the quantum physical decay process were used: 1. A normal binary random sequence (Schmidt's RNG) and 2. a binary Markow-chain (Markow's RNG). The subjects were instructed to shift the light column to the top only by will or wishing. Provisions against fraud and artifacts are discussed.

The following psychological variables of the subjects were measured by questionnaires before the experiment: Sheep-goat (SG), locus of control (IPC), personality traits (FPI), actual state of mood (EWL) and the assessment of the subject regarding his or her success in the experiment (confidence score). The physical variables consist of 18 different statistical test quantities which describe different statistical features and reflect theoretical and empirical models derived from the pilot study.

Every subject performed 8 runs consisting of 600 trials each under 4 different conditions in a double blind, respective blind setting

(computer controlled): Markow's RNG with and without feedback and the same for the Schmidt's RNG.

Partly the physical variables consist of a set of statistical test quantities which indicate different features of non-randomness. Others test rather distinct models which were formulated on the basis of the pilot experiment. A further set of physical variables was especially designed to operationalize the concept of pragmatic information. The purpose of these physical variables is not only to fulfil the requirements of the Brunswick's lens model but also to find a specific physical variable (tracer) which allows to distinguish 'PK-signals' from the 'noise' of the random source.

The fundamental hypotheses of the experiment are called: PK hypothesis, feedback hypothesis, generator hypothesis, tracer hypothesis and the hypothesis of pragmatic information. They are discussed on a more general level.

EUR.J.PARAPSYCHOL., V.6, 1987

A MULTIVARIATE PK EXPERIMENT. PART II. RELATIONSHIPS BETWEEN PSYCHOLOGICAL VARIABLES

Walter von Lucadou, University of Utrecht
Barbara Lay, Universitaet Freiburg
Heidi Kunzmann, Kassel

In part I the general outline of a multivariate PK experiment is described and details of the experimental design are presented. It was emphasized that PK should be demonstrated as a correlation between a set of psychological and a set of physical data. The physical variables and the corresponding general hypotheses were discussed. In the present paper we will discuss the psychological variables and their relationship among each other. Furthermore we will formulate the hypotheses as regards the psychological variables and their relationship to the PK experiment.

2. HOW TO INCORPORATE PSI-EXPERIMENTS IN A NORMAL RESEARCH PROGRAMME?

The purpose for including also questions from ´normal´ psychology is twofold:

(i.) In general, we want to emphasize that parapsychology and ´normal´ psychology should not be regarded as antagonistic fields since both contribute to the scientific exploration of the human psyche or human behaviour. Furthermore, in accordance with Irwin´s

(Irwin 1979) model, it is assumed that 'normal' human abilities and 'psychic' abilities may depend on similar or even the same psychological factors or categories. This idea implies that it would not be wise to 'isolate' psi too much from normal psychology.
(ii.) In process oriented parapsychological experiments, there is a possibility that either no psi effect can be demonstrated, which makes it impossible to test the experimental hypotheses, or that the results are so confusing that one cannot even apply the hypotheses formulated beforehand. As Houtkooper (1983) stated, his major finding seemed to be inconsistency: 'In my own research, inconsistent results are a major landmark. Straight replication has been a good recipe for significant results to change sign!' (p.87). This elusiveness of psi normally puts a lot of pressure on the experimenter, especially in the case of expensive and time consuming experiments. Under the present conditions at the (German) Universities no experimenter can afford to spend a lot of time and effort on uncertain experiments. This becomes even more difficult because of the sometimes hostile atmosphere parapsychology has to tangle with. It is therefore an important part of the experimental design of a parapsychological experiment to reduce the pressure on the experimenter. This should not only be seen in the light of conscious or unconscious data manipulations (see Broad & Wade 1984) or a possible 'experimenter effect' (see Rosenthal 1969) but also from the point of view of the psychological setting. It may well be that a tense atmosphere during a (para-) psychological experiment may distort the results. Such tensions can at least be reduced if 'normal' psychological questions can be included in such a way that they do not only guarantee a recognized scientific aspect but can also be used to justify the experimental setting of the psi-experiment.

In the PK experiment described here, the question incorporated from 'normal' psychology was whether personality characteristics or belief-systems may distort the judgement or assessment capacities of persons who attribute own 'abilities' to certain fluctuations or outcomes of the random process. The assessment of uncertain physical processes by persons is a subject of mainstream psychology which gains more and more interest among psychologists (see Lichtenstein et al. 1980). Another question from 'normal' psychology is the interdependence of different personality concepts (traits) and state variables to the sheep-goat concept developed in parapsychology. Furthermore, the intercorrelations between traits and state variables may be of interest. Such investigations may shed some light on the psychological structure of belief-systems. It may be interesting to note that those psychologists who are sceptical about parapsychology usually emphasize the necessity of research in ´occult´ or ´superstitious´ belief-systems (see Alcock 1981, Spada 1987).

3. PSYCHOLOGICAL QUESTIONS IN PK RESEARCH

There are, of course, also a lot of psychological questions with are traditionally investigated in psi experiments such as the sheep-goat hypothesis or the dependency of ´psi abilities´ on personality constructs. However, most of this research has been done in ESP research. Only a few systematic investigations on psychological variables in PK experiments are available, and most of these do not exhibit clear relationships (see Dale 1946, Nash 1946, Nash & Richards 1947, Van de Castle 1958, Mischo & Weis 1973).

Based on more or less qualitative empirical data in PK-research some psychological variables seem to be of special interest. Thus Roll (1985) expects that ´masculinity´ might be a predominant predictor for success in PK experiments, but not in ESP-experiments. Such a difference can also be expected with the sheep-goat variable. For this variable a rather reliable correlation could be found for ESP-tasks but not for PK. However, a main difficulty for the validation of such findings is the problem of comparability, or concept validity, and reliability of the psychological descriptors.

Theoretical considerations or models may provide another approach to specify the psychological conditions of PK-tasks. In this respect the conformance behaviour model of Stanford (1974) or Batcheldor´s (1979) model of ´ownership resistance´ and ´witness inhibition´ might be useful. One should be aware, however, that these two models use different psychological concepts about PK. Stanford´s model describes PK mainly as a ´disposition´ depending on situational factors which are necessary to trigger a psi-effect, whereas Batcheldor´s model primarily adopts the concept of a ´PK-ability´ which, for instance, may be (sub-) consciously suppressed. Moreover, Stanford´s model primarily describes spontaneous psi events whilst Batcheldor´s model was developed from sitter-group experiments. However, both models may be used to derive hypotheses for formal statistical PK experiments. In contrast, other models would deny such a close relationship between PK and personality characteristics because they emphasize the unique (archetypal) character of psi events in the history of life of a person. As an example may serve C.G. Jung´s idea of synchronicity (see

v. Franz 1983).

4. THE PSYCHOLOGICAL VARIABLES

In the present study a set of 4 psychological questionnaires were used to measure the independent psychological variables. In addition confidence-scores of the subjects were measured before and after each run. The confidence-scores can be regarded either as dependent variables, for instance, in relation to the study of judgement characteristics, or as independent variables in relation to the PK-task.

Thus, the raw data on psychological variables consist of the following five sets:
(i.) A sheep-goat questionnaire (SG) consisting of 18 items which was especially developed for this study.
(ii.) The IPC-questionnaire (Krampen 1981) measuring locus of control (according to Rotter 1966) which includes 24 items.
(iii.) A short version of the ´Freiburg Personality Inventory (FPI-K) including 76 items and 4 social items (Fahrenberg et al. 1978).
(iv.) A shortened version of a mood scale (list of adjectives) EWL-K consisting of 66 adjectives and 4 items concerning the feasibility of the test (Janke and Debus 1978).
(v.) The confidence-score before and after each run (8 items) in which the subject expresses whether he or she feels confident to ´produce PK´ during the following run and whether he or she believes that the previous run was successful. Confidence-scores are expressed on a scale from 1 to 5.

The statistical evaluation of these psychological raw data is described in detail in the thesis of one of the authors (Lay 1982). Here, for brevity, we will only summarize the results.

For the SG-, IPC- and the EWL-questionnaires a factoranalysis was carried out. Since the SG-questionnaire was composed of several previously developed, extensive SG-questionnaires (see Huber 1979, Ballstaedt 1979, Thalbourne & Haraldsson 1980 and Thalbourne 1981) it was necessary to apply an item-analysis with subsequent factor analysis. A factor analysis was also necessary for the EWL-questionnaire, because we used a shortened version instead of the original one. In the case of the IPC-questionnaire the original sample (n=151) was much smaller than ours. Only for the FPI-K we decided to use the standard scales for further evaluation since this version is a well established tool as regards its reliability and validity. Moreover, the FPI-K was also used in a previous PK experiment (Weis 1972, Mischo & Weis 1973) which allows a direct comparison.

The factoranalyses of these questionnaires, however, revealed that only the SG showed a clear factor structure. Thus one should not overestimate the discriminatory power of the factors of the IPC and the EWL. The absence of a clear factor structure could be interpreted as the consequence of a ´reluctant´ test attitude of the subjects who might not have understood the relevance of rather personal items in relation to a PK experiment (see Lay 1982, 6.2.5). Or perhaps the sample of subjects was strongly biased concerning these psychological traits.

The results of the SG-questionnaire are of special interest: The itemanalysis of the 18 items showed that consent of the ´sheep´ was less distinct than the refusal of the ´goats´ though both groups turned out to be rather balanced. The mean value of the total score was x = 36.20 and the standard deviation SD = 6.76 with a range of 22 to 54 in relation to the expectation value of x = 36 and a range of 18 to 54. The estimation of the inner consistency with Cronbach´s = .83 indicates the reliability of the item pool.

As criteria to exclude items a internal correlation coefficient rit (selectivity coefficient) and the skewness of the distribution x (difficulty index) was used. It turned out that 3 items had to be omitted. Item 12 due to unsufficient selectivity (rit = .28), item 9 due to extreme skewness (x = 2.87), and item 17 also due to extreme skewness (x = 2.86). The subsequent factor analysis revealed a three factor solution meeting the eigenvalue criterion and explaining 55.3 % of the total variance. The VARIMAX rotated, orthogonal factormatrix can easily be interpreted:

(i.) SG-factor 1 (7 items, 33.6% of the total variance): Personal experiences of ESP, especially precognition.
(ii.) SG-factor 2 (7 items, 12.4%): Belief in PK as a scientific proven fact, mainly concerning items about experimental and theoretical possibility of PK.
(iii.) SG-factor 3 (3 items, 9.3%): Confidence to demonstrate personal ESP- or PK-abilities under laboratory conditions.

For the IPC-questionnaire also three factors were found, however, the differences are less distinct:

(i.) IPC-factor 1 (13 items, 17.3%): Externality. Feeling of subjective powerlessness. The own life and behaviour is experienced as dependent from other powerful persons.
(ii.) IPC-factor 2 (10 items, 10.1%): Internality. Events in the personal environment and the own life history are mainly considered to be under the persons own control.
(iii.) IPC-factor 3 (4 items, 7.3%): Fatalism concerning car accidents. Refusal to accept one´s own responsibility.

The data of the EWL-questionnaire exhibited a five-factor structure concerning aspects of present mood:

(i.) EWL-factor 1 (25 items, 24.3%): General desactivity, tiredness, numbness.
(ii.) EWL-factor 2 (25 items, 8.8%): Depressed mood, irritability.
(iii.) EWL-factor 3 (18 Items, 4.6%): High spirits, activity, cheerfulness, self-confidence.
(iv.) EWL-factor 4 (13 items, 4.0%): Nervousness, excitedness, anxiousness.
(v.) EWL-factor 5 (7 items, 3.5%): Frustration, rage.

The 12 standard scales of the FPI-K can be named as follows:

(i.) FPI-scale 1: nervousness.
(ii.) FPI-scale 2: spontaneous aggressiveness.
(iii.) FPI-scale 3: depressivity.
(iv.) FPI-scale 4: excitability.
(v.) FPI-scale 5: sociability.
(vi.) FPI-scale 6: calmness.
(vii.) FPI-scale 7: dominance.
(iix.) FPI-scale 8: inhibition.
(ix.) FPI-scale 9: openness.
(x.) FPI-scale 10: extraversion.
(xi.) FPI-scale 11: lability.
(xii.) FPI-scale 12: masculinity.

It is important to note that the names of these psychological variables do not describe their meaning precisely. The names are merely labels which are used for convenience. For a closer psychological interpretation it would be necessary to go back to the factor matrices which describe the weight of each item on each factor. The meaning of each factor is derived from the meaning of its corresponding items. (A detailed description of the meaning of the variables can also be found in Fahrenberg, et al. 1978 and Lay 1982).

The assessment of the subject's 'disposition' before each feedback-run (1,3,5,7) (confidence-score: PRIOR1 - PRIOR7) and the assessment of success after each feedback-run (POST1 - POST7) was based on the single responses. It is important to note that the values are negatively poled: 1 means a very positive assessment while 5 means a very negative one. The following table presents the mean value x, the standard deviation s and the deviation u from the expectation value x0 (with x0 = 3; u = (x - x0)/(s*sqrt(n))) for each item.

TABLE 1
The confidence-score distributions

Confidence-score before run	x	s	u	
1	3.037	1.078	0.59	
3	2.893	1.139	-1.62	
5	2.977	1.246	-0.32	
7	3.010	1.332	0.13	
after run	**x**	**s**	**u**	
1	2.773	1.063	-3.69	**
3	3.231	1.166	3.43	**
5	3.110	1.169	1.64	
7	3.241	1.205	3.46	**

It is interesting to compare the assessment of the subjects before and after each run for successive trials of the session. After having given their confidence-score concerning the success they expect they received feedback on their real achievement. Thus the change in this score during the progression of the session might reflect the subjects ability to learn to interpret the trial by trial feedback during the previous runs.

From the final evaluation of the confidence-score it appeared that before the first run the subjects tend to give a rather cautious (slightly negative) assessment. However after the first run they show a highly significant positive assessment about their success, which may indicate that they misinterpret the feedback during the run. Moreover, this positive assessment seems to continue in spite of the realistic feedback during the pause, because the assessment before the second run is again positive (but not statistically significant). However, after the second feedback-run (run 3) the assessment of success becomes significantly negative and remains negative while the confidence-score before the third feedback-run (run 5) still remains slightly positive before it returns to a rather neutral value before the last run (run 7). This result indeed indicates that, concerning the statistical mean values, the subjects are able to adapt their assessment due to the feedback during the pauses.

5. RELATIONSHIP BETWEEN THE PSYCHOLOGICAL VARIABLES

It is obvious that a number of n psychological variables (n=31, exclusive the social data) would theoretically allow an amount of n*(n-1) (=930) possible combinations of variables whose interdependencies could be investigated. However, the underlying factor model of personality (see Cattell 1950, Eysenck 1953) implies that most of such possible combinations must be meaningless. The method of factor analysis involves that the correlations among the factors of one questionnaire are minimalized. Therefore it is not useful to study correlation between factors of the same questionnaire. However, in the case standard scales are used instead of the factors this might be questionable, but only if the questionnaires are not reliable. The same should hold for different personality traits. In the context of the factor model of personality it is a requirement that valid personality variables should be rather independent among each other. For this reason neither the intercorrelations of the FPI-scales, nor those of other traits were investigated.

A similar argument can be used for the trait - state distinction.

Traits are those variables which have been shown to be rather stable during a longer period of life of a person and which do not depend on the present state. On the other hand it is likely that the momentary mood (state) of a person depends on his or her personality characteristics (traits) as well as on situational factors. Since the situational factors cannot be controlled over a sufficient period of time before the experiment it is rather unlikely that possible correlations between these sets of variables contain information which can be meaningfully interpreted in the context of the experiment.

As regards of the sheep-goat variable it is still unclear whether it should be regarded as a state or as a trait variable. Its intercorrelations to other personality variables are not known in detail. Hence it is not clear whether the sheep-goat variable is indeed an independent and valid construct or just another name for an already known trait. Since our confidence-scores could also be interpreted as a sheep-goat state variable, we are able to contribute to the sheep-goat research on the basis of our data. Several recent investigations, however, seem to indicate that the sheep-goat variable can be considered as a valid trait variable (for a recent overview see Mischo 1982, Mischo & Lay 1985).

Thus the question of the relationship among the psychological variables turns out to be mainly the question of the relationship between the SG-variables, respectively the confidence-score, and the other psychological variables.

The investigation of the relationship between the SG-variables and the other psychological variables is described in detail in the thesis of one of the authors (Kunzmann 1984). Here we will only report the main results.

The intercorrelations of the different psychological variables were investigated by three statistical methods: (1) Calculation of correlation coefficients. It turned out that Pearson´s product-moment correlation coefficients showed more ore less the same dependencies as Spearman´s rank correlation coefficients. (2) Simple analysis of variance to test for differences of the mean values of different groups. (3) Comparison of the mean values of extreme groups by t-tests.

For the analysis of variance and the t-tests the predictor variables had to be divided in different classes or groups. For the FPI-scales the 4 groups emerged directly from the given scales (quartiles). In a similar way, the five possible values of the confidence-scores were used as classification of the groups. For the factors of the SG-, IPC- and EWL-questionnaires the corresponding distributions of the total score was divided into 5 groups with about equal group size.

Table 2 presents the relationships between the different psychological variables. The upper part of the table gives the significant Spearman's rank correlation coefficients and the lower part the p-values of the analyses of variance who turned out to be significant. The p-values of the significant t-test analyses are given in parentheses. The significance level was set at $p = .05$. In the case of the analyses of variance and t-tests the variables at the left hand side of the table are considered to be the independent variables. In addition to the confidence-score before and after each feedback-run the sums of the four corresponding scores are also given (PIORS and POSTS). As already stated, no correlation coefficients or p-values were calculated for the relationships among the FPI, IPC and EWL variables.

The data from table 2 indicate that the relationship between the SG-factors and the other questionnaires is rather poor. There are only relationships between SG-factor 1 and the FPI-scales 5, 6, 8, and 10, and the IPC-factor 3. However, the different statistical methods complement each other in a meaningful way. Therefore the following conclusions can be considered as statistically valid (for further details see Kunzmann 1984):

a) The relationship between the SG-variables and the other psychological variables IPC, FPI, EWL seems to be rather weak. Persons who have a high score on the variables sociability, calmness, and extraversion more often describe own personal experiences of ESP compared to persons with a low score on these variables.

b) Less plausible seems the result that persons with a high score in the IPC factor 3 also frequently report personal ESP experiences. IPC-factor 3 describes persons with a fatalistic attitude concerning car accidents and who reject responsibility in this respect.

According to the factor model these findings suggest that the SG-variable can be regarded as an independent variable. This result is in agreement with the findings of Thalbourne & Haraldsson (1980). Especially the two factors concerning PK (SG-factor 2 and SG-factor 3)

TABLE 2

Relationship between the psychological variables

	PRIOR					POST					SGFAK			IPCFAK		FPISK										EWLFAK				
	1	3	5	7	S	1	3	5	7	S	1	2	3	1	3	1	3	4	5	6	7	8	10	11	12	1	2	3	4	5
PRIOR1	--	.53	.41	.47		.28	.31	.22	.31		-.18		-.17			.14				-.13		.15						-.16		
PRIOR3		--	.42	.53		.41	.42	.39	.37		-.20		-.15																	
PRIOR5			--	.52		.20	.37	.38	.35		-.17		-.13			.14				-.14	.13									
PRIOR7				--		.21	.30	.42	.45		-.18	-.17				.16		.13												
PRIORS					--					.63	-.24	-.14	-.17			.18				-.16	.14									
POST1						--	.27	.26	.31																					-.13
POST3							--	.20	.40		-.19											.13				-.15				
POST5								--	.35																					
POST7									--		-.16																			
POSTS										--	-.21									-.14								-.13		
SGFAK1	.018	.008	.019	.016	.001		.014			.012	--				**				.14	.15		.13	.21							
SGFAK2			.022	.020	.019					.047		--																		
SGFAK3	.005	.052	.034	.042	.003					.043			--																	
IPCFK1														--																
IPCFK3											.022				--															
FPISK1	.039		.018	.008	.004											--														
FPISK3								.024									--													
FPISK4											.044							--												
FPISK5											.029								--											
FPISK6			.054								.011									--										
FPISK7																					--									
FPISK8	.049				.046		.015															--								
FPIS10		.052	.042		.020						.001												--							
FPIS11									.052															--						
FPIS12			.043																						--					
EWLFK1	.050				.047		.034																			--				
EWLFK2	.010		.030	.003	.007					.046																	--			
EWLFK3	.043																											--		
EWLFK4	.046																												--	
EWLFK5						.014																								--

show no significant correlation to any of the psychological variables. This indicates that the belief in PK (or psi), in fact, does not depend on personality factors. Only SG-factor 1 (personal experiences of ESP) seems to be of relevance. However, this factor does not primarily describe a ´belief-system´ but rather personal experiences of ESP. If we assume that such experiences are no illusions but based on real ESP events the positive correlations of FPI-scale 5 (sociability), 6 (calmness) and 10 (extraversion) would fit well with the overall findings in experimental ESP-research (for an overview see Eysenck 1967, Rao 1980, Palmer 1986). Only the positive correlation with FPI-scale 8 (inhibition) does not fit into this picture. This correlation, however, is weaker than the others and is also not supported by the analysis of variance. Hence, these relationships of the SG-variables with the other personality variables could indeed indicate that belief in psi primarily depends on personal paranormal experience.

c) In contrast to the SG (trait-) variables, the relationships between the different confidence-scores (PRIOR1 - 7) or the assessment of the subjects of the feedback during the runs (POST1 - 7) exhibits a rather complicated and manifold structure. Here, we were interested in the question whether and how the PRIOR and POST variables correlate among each other and whether there exist certain tendencies of assessment which exhibit a relationship to personality variables or to the momentary state at the beginning of the experiment. We expected that ´sheepish´, extraverted and/or non-neurotic subjects would tend to assess their success more positive than the corresponding antagonists. We further assumed that subjects in a calm and positive mood would also tend to give positive assessments.

Table 2 shows that the PRIOR and POST variables correlate significantly among each other. This indicates that these variables could be combined into two common factors namely PRIOR and POST. Furthermore, the high correlations between PRIOR and POST show that the assessment of the subject´s success depends on his or her confidence to be successful but that this confidence also depends on the feedback of the previous run (as discussed above). The correlations of the PRIOR variables with the other psychological variables are more numerous than those of the POST variables.

In agreement with our expectations sheep tend to be more confident than goats (remind the negative direction of the PRIOR and POST -scales). In this case it turns out that again the SG-factor 1 (personal ESP experience) plays a predominant role. It is remarkable that the IPC factors (externality, internality and fatalism) do not show any significant correlations to PRIOR or POST.

d) The correlations between the FPI-scales (1, 6 and 8) and the PRIOR variables show more ore less the expected tendencies. The fact that these dependencies seem to fade for the POST variables indicates that the post variables may depend strongly on the feedback.

e) The relationships to the EWL-factors are more difficult to interpret. It seem plausible to assume that the mood of the subjects changed during the experiment. Such a change would result in non-linear dependencies which would wash out the (linear) corrrelations. Indeed it can be seen in table 2 that only 4 significant correlations appear whilst the analysis of variance yields 11 significant relationships.

The results concerning PRIOR and POST assessments can be summarized as follows (Kunzmann 1984, p.88): Those subjects who are very confident in being successful (high score in PRIOR) - in contrast to those who are extremely low scoring - report more frequent ESP experiences, seem to possess more self-confidence, a larger frustration tolerance and less inhibitions or nervous (psychosomatic) complaints. Subjects in a positive and active mood tend to give a more positive confidence-score (PRIOR and POST) than those who do not feel active. In contrast to what would be expected, however, it turned out that subjects who were in nervous and aggressive or angry mood before the experiment tended to be more confident than subjects who were in calm mood.

6. THE PSYCHOLOGICAL HYPOTHESES CONCERNING THE PK-EFFECT

Most of the psychological intercorrelations discussed in the previous paragraph do not only make sense in the context of psychological models but also in relation to parapsychological questions. There seems to be a certain analogy between the ´ability of assessment´ (PRIOR and POST) and a possible ´PK-ability´ of a person. One might feel tempted to transfer the psychological structure detected above also to the ´PK-ability´. However, it is an open question whether PK is indeed an ability.

Keeping in mind this question and assuming that PK could in principle be expressed as a correlation between physical and psychological variables, it is useful to formulate some psychological hypotheses on the basis of the variables measured in this experiment.

Since the sheep-goat hypothesis was so successful in ESP-research it seems justified to apply it also in PK experiments. The confidence-score PRIOR1 could be conceived as a kind of actualized sheep-goat item (or SG-state). However, only PRIOR1 can be used as an independent variable. In PRIOR3, 5 and 7 a ´learning´ effect could lead to artificial (no-PK) correlations! In the light of the IDS-model (May et al. 1985), however, PRIOR1 could also be regarded as an ESP trial. In any case we would assume a positive correlation between ´PK-success´ and the variables SG-factor 1, 2, 3, PRIOR1)

Concerning the IPC-variables it is difficult to formulate a definite hypothesis. Bender et al. (1976) and Sannwald (1962) reported in their analysis of spontaneous cases, that persons with spontaneous PK experiences tend to accept an external locus of control. Therefore we formulated the hypothesis that externality correlates positively with PK-success. It is obvious that from a psychological point of view a negative correlation also seems plausible. Moreover it is questionable whether one can transfer psychological conditions of spontaneous cases to the experimental situation.

In the case of the FPI-variables several hypotheses could be formulated in advance. In a previous multivariate PK experiment Weis (1972) applied the same questionnaire. His results were in general agreement with Eysenck´s (1967) findings but only in the ´frustration´ condition. It is expected that extraverted, calm, sociable and active subjects are more successful in a PK experiment than subjects with high scores on depressivity, neuroticism, nervousness and inhibition.

From the experiment of Weis another hypothesis can be derived. Since he only obtained significant results in the ´frustration´ condition one would assume that the state variable EWL-factor 5 correlates positively to PK-success. In addition, on the basis of Stanford´s model and the qualitative results of the well-known ´Philip-group´ (Owen & Sparrow 1979) one would expect that the state ´high spirits´ (EWL-factor 3) is conductive to PK in contrast to EWL-factors 2 (depressed mood) and 4 (nervousness).

So far the psychological hypotheses are not fully specified, because it is not known what ´PK-success´ means in the context of a PK-correlation. If we describe a PK-effect no longer in terms of an extra-chance deviation of ´hits´ but as a significant correlation we also have to reformulate the psychological hypotheses. Thus the amplitude of an extra chance deviation has to be replaced by the strength of a correlation and the direction of the deviation by the sign of that correlation. In order to compare these values with the psychological hypotheses it is important that we are able to connect the meaning of the psychological hypothesis with the content of the correlation. However, this cannot be done in every case. For instance, if we consider the physical variable ´hit-variance´ (see part I chapter 4.5, 2.) it is not clear whether a positive correlation to a psychological variable means a ´success´ or a ´failure´ of the PK task. Thus the definition of ´PK-success´ in relation to physical variables requires additional knowledge about the process involved.

7. OPERATIONALIZATION OF PRAGMATIC INFORMATION

Obviously, the term ´success´ in a PK experiment cannot be described in absolute terms. It depends not only on the instruction given to the subject but also on the experimental situation which creates the context for the meaning of ´success´ for both subject and experimenter. However, in regard to the psychological variables only the meaning of success as experienced by the subject is of relevance. This meaning mainly depends on the feedback information provided to the subject.

In part I, paragraph 7 and in a recent paper (v. Lucadou 1987) it is argued that the meaning of the term ´success´ can be best described by the concept of pragmatic information. Here we describe the specific approach of operationalization we applied in this experiment. This does not mean that other more successful operationalizations would not be possible, but it provides an example of how to proceed and it shows that any operationalization has to start from certain psychological assumptions or models.

From these general considerations follows that the variables V, R, D, K, K´, ZTA, ZB, ZC, ZD, and ZE (see Part I, paragraph 4.5) cannot be used to describe the pragmatic information of the random sequences for the subject because they are not related to feedback or instruction. It is important to note that any tracer (see part I, 7.)

would by definition not be suitable for a measure of pragmatic information because a tracer is defined as an indication of a PK-effect without making reference to psychological conditions. The number of hits T might be used as a measure for pragmatic information, because the term ´hit´ is closely related with the instruction. In our case, however, the number of hits T is not related to the display information in a simple one to one correspondence. As already described in Part I, 4.1, the subject cannot detect further ´hits´ if the light column has reached the top of the row and respectively no further ´miss´ if the column has reached the bottom. The ´display function´ is best represented by the variable DIM (Part I, 4.5, 18.). Hence the variable DIM may be regarded as the best measure for the pragmatic information of the feedback information.

From theoretical considerations (see v. Lucadou 1987, v. Lucadou & Kornwachs 1983) a more detailed structure of the measure of pragmatic information was proposed. Pragmatic information is conceived as a product of two independent complementary concepts: ´novelty´ (Erstmaligkeit) and ´confirmation´ (Bestaetigung) (see v. Weizsaecker 1973). Thus it would be interesting to try an operationalization of these two concepts in the context of our experiment.

From a psychological point of view it is plausible to assume that a subject observing the display would experience any sudden rise of the light column as a ´success´ but also as something new, namely ´novelty´. A sudden fall would also be experienced as ´novelty´ but as ´failure´. Thus we can distinguish two different types of novelty: One attributed with positive fluctuations of the random sequence and one with negative fluctuations. Positive fluctuations are measured by the variable ETS (part I, 4.5, 12.), negative fluctuations by the variable ETG (15.)

If a fluctuation would last a certain time interval it is plausible to assume that this would be experienced by the subject as ´confirmation´ of his or her success or failure. Hence the duration of a fluctuation (variables BES (13.) and BEG (16.)) may serve as operationalization of the term confirmation. Again positive and negative fluctuations have to be distinguished because of their different meaning for the subject.

Since pragmatic information is assumed to be the product of novelty and confirmation the variables PIS (14.) and PIG (17.), which are based on products of values of ETS and BES or ETG and BEG for each trial, can be considered as an operationalization of the pragmatic information concerning 'success' and 'failure' of the subject. At first glance it may seem incomprehensible that the two different concepts 'success' and 'failure' are used independently, in contrast to the variable DIM which has a one-dimensional structure. However, this becomes more clear if one consider that some subjects might feel both successful and unsuccessful during a run whilst others might experience only one or none of these possibilities. This means that in the context of this operationalization the pragmatic information of success and failure are independent variables. A factor analysis of all physical variables (which will be discussed in part III), indeed showed that the variables PIS and PIG are associated with independent factors.

It is obvious from our discussion of the psychological variables in the previous paragraph that the pragmatic information of success and failure corresponds with the sheep-goat variable. Thus we would expect that sheep will obtain high scores in PIS in contrast to goats who will get high scores in PIG. However, since there are three different sheep-goat factors it is not quite clear which combinations can be expected. In any case, in terms of correlations one could say that the sheep-goat variables should exhibit stronger correlation to the variables PIS and PIG than to the variable T or to variables which do not contain pragmatic feedback information. Furthermore, the correlations between SG and PIS and between SG and PIG should show opposite signs, namely PIS a positive and PIG a negative one. Concerning variable DIM one would also expect a strong correlation to the SG variables, but it is difficult to predict a priori the proportion of the size of the corresponding correlation coefficients.

Since most psychological variables are defined on one-dimensional scales in relation to their possible dependency on the 'PK-success', it might be possible that the distinction of two independent measures PIS and PIG turns out not to be useful. Correlations between bipolar and unipolar variables could lead to a diminished magnitude of the corresponding correlation coefficients. However, this could easily be detected by comparing with the corresponding correlation coefficients of the variable DIM, which involves a bipolar operationalization of the pragmatic information of 'PK-success'.

Unfortunately, a closer analysis of the variables ETS and BES respectively ETG and BEG revealed that the chosen operationalization did not work as was expected. It is an important theoretical requirement that the measures for novelty and confirmation should be independent. This is (among others) a special condition which has to be fulfilled for complementary concepts. Table 3, however, shows that ETS and BES as well as ETG and BEG are highly correlated for both Markow- and Schmidt sequences. Consequently these variables also correlate with PIS and PIG respectively.

TABLE 3
Correlations between the components of pragmatic information

Variable pair	Markow run	Schmidt run
ETS/BES	.84	.95
ETG/BEG	.85	.97
ETS/PIS	.84	.85
ETG/PIG	.88	.89
BES/PIS	.79	.86
BEG/PIG	.85	.89

This unfortunate result does not necessarily imply that the idea of this operationalization has to be given up. It seems plausible that the correlations in table 3 are produced by a special condition in the operationalization of ETS and ETG. As described in part I (4.5, 12.) only those fluctuations were counted which exceed a certain threshold. This threshold (CR(n) = 2), however, was not fixed during the run because it is dependent on the length of a ´dynamic´ interval. This interval was the number of steps from the present position of the light column on the display to the top of the row. It might have been better to use a fixed threshold to possibly avoid a correlation between ETS and BEG.

Such alternative operationalizations could be tried in future experiments. In this study, however, it was decided beforehand that neither the evaluation procedures nor the hypotheses were to be changed on the basis of experimental results. For this reason the original formulation of our hypothesis concerning the psychological aspect of the PK experiment is maintained.

8. FORMULATION OF THE PSYCHOLOGICAL PK HYPOTHESES

In contrast to paragraph 7 of part I the hypotheses for the main experiment will be formulated here in such a way that they refer to test statistics which are used in the evaluation. As described in more detail in part III two methods were used to detect possible relationships between the psychological (as independent) and physical (as dependent) variables. These two methods are rank variance analyses and rank correlations. For the following formulation of the hypotheses no distinction between these two methods is made. A relationship between two variables may be exhibited by only one or by both of the two methods. For the validation of the relationship, however, this has to be taken into account. Any such relationship will formally be described by the symbol ri,j(v,w). The symbols i and j describe different psychological variables (factors) of a specified questionnaire which may show a relationship r to a physical variable v or w. (Since the scales of DIM and PRIOR are negatively poled this will be indicated by -DIM and -PRIOR). The hypotheses are described here one after another, but it should be kept in mind that according to figure 8 in part I they are not independent among each other.

1. The PK hypothesis (1)
H1: There exists at least one 'meaningful' (in respect to the other hypotheses) and significant relationship r between psychological and physical variables.
H0: No significant 'meaningful' r can be found.

2. The feedback hypothesis (2)
H1: There exists a significant difference between feedback- and

non-feedback conditions (especially for the variables T and DIM).
H0: No significant difference can be found.

3. The sheep-goat hypothesis (3)
H1: At least one of the sheep-goat variables is positively related to T or DIM:

$r_{SG,-PRIOR(-DIM,T)} > 0$

H0: No significant r can be found or r is negative.

4. The externality hypothesis (4)
H1: At least one of the following relationships is significant:

$r_{IPC1,3(-DIM,T)} > 0$

$r_{IPC2(-DIM,T)} < 0$

H0: No significant r can be found or all are reversely poled.

5. The extraversion hypothesis (5)
H1: At least one of the following relationships is significant:

$r_{FPI5,6,10,12(-DIM,T)} > 0$

$r_{FPI1,3,4,8(-DIM,T)} < 0$

H0: No significant r can be found or all are reversely poled.

6. The frustration hypothesis (6)
H1: Frustration correlates positively to PK-success.

$r_{EWL5(-DIM,T)} > 0$

H0: r is negative or insignificant.

7. The hypothesis of high spirits (7)
H1: Positive mood correlates positively to PK-success.

$r_{EWL3(-DIM,T)} > 0$

$r_{EWL1,2(-DIM,T)} < 0$

H0: No significant r can be found or all are reversely poled.

8. The aggregation hypothesis (8)
H1: Aggregations of variables according to Brunswick's lens model show an increased r between psychological and physical variables.

$r_{i(\sum v)} > r_{i(v)}$

$r_{\sum i(v)} > r_{i(v)}$

H0: r remains equal or decreases.

9. The hypothesis of pragmatic information of the display (19)
H1: The variable DIM shows stronger relationships to all psychological variables than the variable T.

$$ri(-DIM) > ri(T)$$

H0: $ri(T) \leq ri(-DIM)$

10. The hypothesis of pragmatic information on success and failure (20)

H1: The variables PIS and PIG show a stronger relationship r to the sheep-goat-variables than the variable T in hypothesis 3 and the correlations for PIS and PIG will have opposite signs.

$$rSG(PIS) > rSG(T)$$
$$rSG(PIG) > rSG(T)$$
$$rSG(PIS) > 0$$
$$rSG(PIG) < 0$$

H0: The variables PIS and PIG show weaker relationships than T or the signs are in the wrong direction.

In part III the remaining hypotheses concerning the physical aspects of the PK experiment are discussed. Furthermore the methods of evaluation will be described and the results of both pilot and main experiment will be reported and discussed.

ABSTRACT

This paper is the second of three parts on a multivariate PK-study. It mainly describes the psychological variables and hypotheses.

It is shown that a PK experiment can be designed in such a way that it also covers research aspects from ´normal´ mainstream psychology. This approach not only indicates that parapsychology should be considered as a part of ´normal´ psychology but should also reduce the pressure on experimenters working in anomalies research. Such pressure could also distort the psychological setting of a PK experiment.

The psychological variables were measured with four questionnaires: A newly designed sheep-goat questionnaire, the IPC-questionnaire (locus-of control), a personality-questionnaire (FPI-K) and a mood scale (EWL). Only for the FPI-K the original scales were used as variables. The other questionnaires were subjected to a factor analysis. The resulting factors were used as independent variables. An additional confidence-score describing the subject´s assessment of his or her PK-success in the following respectively the previous run was used in two different ways: (i) As dependent variable to describe the

subjects judgement abilities in relation to personality traits, present states and belief-system (SG) and (ii) as independent variable (SG-state) as a predictor of PK-success.

The evaluation of the psychological data revealed (i) that the sheep-goat construct can be regarded as a valid variable, independent from other personality variables, except from SG-factor 1 which (ii) indicates that personal experience of ESP seems to rely on genuine ESP rather than on belief systems. In contrast, the data on the confidence-score show a lot of correlations to trait and state variables according to the psychological expectations. Furthermore the data support the assumption that the subjects are able to adjust their assessment according to the feedback information.

In addition to theoretical considerations and previous empirical findings the assumption that 'PK-abilities' exhibit a similar psychological structure as the relationship between the confidence and the trait variables leads to the formulation of the psychological hypotheses of the PK experiment.

Special hypotheses are formulated on the basis of a specific operationalization of the concept of pragmatic information. However, it turned out that in one case the operationalization did not meet the theoretical requirements. Nevertheless the formulation of the hypothesis was retained in its original form.

EUR.J.PARAPSYCHOL., V.6, 1987

A MULTIVARIATE PK EXPERIMENT. PART III. IS PK A REAL FORCE? THE RESULTS AND THEIR INTERPRETATION

Walter von Lucadou
University of Utrecht

In part I the fundamental hypotheses of the multivariate PK-experiment were discussed and the experimental setting described. In part II the psychological aspects of the experiment were reported. On the assumption that ´PK-abilities´ exhibit a similar psychological structure as ´assessment abilities´ a precise formulation of the psychological hypotheses of the PK experiment could be given. Since in the context of this experiment PK is only described as a correlation between independent psychological and physical variables it is necessary to reformulate the general PK hypotheses of part I in terms of correlations between the psychological variables described in part II and the physical variables described in part I. In this paper the evaluation techniques are described before the results of the experiment are reported and discussed.

2. THE PHYSICAL HYPOTHESES

From figure 8 (of part I) it can be seen that the generator hypothesis, the accumulation hypothesis, the tracer hypotheses, and the influence hypothesis (part I, paragraph 7), can be considered as the main physical hypotheses. However, they cannot be tested directly but their validation depends on the specific formulation (operationalization) of the physical variables and hence on more detailed hypotheses. These more detailed physical hypotheses are indicated in figure 8 (of part I) by a dark hatching. Thus the tracer hypothesis (for instance) can be investigated not only by the variable K as will be described below but also by any physical variable 1 - 18 (paragraph 4.5, part I). In the case a tracer is found one would expect that the global distribution of the corresponding variable shows a significant and systematic deviation from its 'normal' distribution under the null hypothesis. This implies that the physical hypotheses related to a tracer effect are not totally independent. On the other hand the physical hypotheses concern mainly the physical aspects of PK or of the alleged 'PK-mechanism' rather than specific psychological relationships. Therefore it is important to keep in mind that the physical hypotheses can only be evaluated if the PK hypothesis (1), operationalized as a correlation between physical and psychological variables, is first verified. (Numbers given after hypotheses refer to the numbers in figure 8 of part I). This simply means that the PK-process can only be investigated if we know that PK was really present.

1. The influence hypothesis (9) which assumes that PK is a specific influence of the subject on the radio-active source is mainly operationalized by the cross-correlation test quantity K' (6). (Numbers after quantities refer to paragraph 4.5 of part I in which descriptions of these quantities are presented). A significant value of K' indicates a simultaneous signal on all five counters. However, to be general enough to reject the null hypothesis it is only required that at least two of these five counters (including the display channel) show a common influence (which means a small value of K').

H1: At least for the 'successful' PK-runs one (i) of the four cross-correlation test quantities (Ki') is consistently one standard deviation or more smaller than the expectation value.
H0: The values of Ki' are increased or fluctuate inconsistently.

2. The tracer hypothesis (10) is a generalization of the influence hypothesis and assumes that the PK effect can be detected by a 'tracer' or a typical process variable which allows to select PK-runs without making reference to the psychological data. The specific operationalization we used here is given by a scattergram of the quantity T versus the auto-correlation test quantity K (5. part I,

4.5) (figure 7 of part I). Unfortunately it is as yet not possible to quantify the accuracy and generality of this operationalization. This means that we do not know which kinds of PK-signals cannot be detected by the method (generality) and what the minimum amplitude of a detectable signal would be (accuracy). The answer to this question can only be found by simulation studies.

H1: At least for the ´successful´ PK-runs the points (T/K) lay (by at least about one standard deviation distance) outside the curve which is produced by random sequences.
H0: All points (T/K) hit the curve.

3. The generator hypothesis (11) is not specified by a single variable since any unspecified (but significant) difference of a PK-effect between the two types of RNGs would reject the hypothesis. The hypothesis tests a fundamental assumption of Schmidt´s model namely that the PK-effect should only depend on psychological and situative variables but not on the ´complexity´ of the target generation and thus not on the type of the RNG.

H1: There is no significant difference between the correlations of the psychological variables to the physical variables of the Markow - and the Schmidt type RNG.
H0: There are significant differences.

Such differences could concern the magnitude as well as the direction of a correlation. However, it should be kept in mind that some physical variables (see Part I, 4.5) of the Markow- and the Schmidt sequences cannot be compared directly. Thus the formulation refers only to comparable variables (such as T, V or DIM).

The remaining hypotheses are more specific and involve specific assumptions about a possible PK process. The verbal description of the hypotheses, however, may lack of precision. In any case the underlying model is defined be the formulation of H1/H0 and not primarily by the given interpretation.

4. The accumulation hypothesis (12) tests the question whether the PK-effect depends on the number of incoming random pulses which are sampled to produce a single trial. If PK would be a process which can produce additional random pulses of the radio-active source, one could expect that a longer sampling time would increase the PK-effect. Unfortunately it turned out, that we were not able to evaluate our

material on the basis of single runs. Therefore we could not test this hypothesis based on correlations but only by using the hit rate.

H1: The hit rate increases with the sampling time.
H0: The hit rate decreases or does not change or changes inconsistently with the sampling time.

5. The variance hypothesis (13) is based on the idea that PK is a process which cannot be controlled voluntarily by the subject. Hence it can be assumed that the PK-performance fluctuates in time. In this case one would expect the PK-effect to show up in the variance or runscore variance of the run rather than in the hit rate. This means that the correlations of V (variance) and R (run-score variance) to the psychological variables should be stronger than for the hit rate T:

H1: The test values V and R show stronger correlations to certain psychological variables than the test value T.

$$ri(V) > ri(T) \quad \text{or}$$
$$ri(R) > ri(T)$$

H0: No psychological variable shows a stronger correlation to V or R than to T.

This weak formulation of the hypothesis is used because it should be avoided that the hypothesis is rejected in the case that only very distinct psychological variables show such a correlation to R and V. It may well be that V, R and T show high correlations to psychological variables but to different ones. In this case the model would still be right in saying that V and R are psychologically relevant or sensitive variables (see paragraph 7).

6. A similar idea leads to the distribution hypothesis (14). This hypothesis is related to the question whether the shape of the distribution of the hit rate may contain more information about the PK-process than the hit rate itself. For instance, one could assume that subjects produce fluctuations in the hit rate which do not affect the mean value but only the shape of the distribution. Both the auto-correlation test quantity K (5) and the Kolmogorow-Smirnov-test quantity D (4) indicate such deviations from an a priori distribution.

H1: The test values K and D show stronger correlations to certain psychological variables than the test value T.

ri(K) > ri(T) or
ri(D) > ri(T)

H0: No psychological variable shows a stronger correlation to K or D than to T.

7. The hypothesis concerning the target generation (time) (15) is closely linked to the specific way of target generation applied in this experiment. As already discussed in part I, the variable ZTA (7) measures mainly short fluctuations of the generation-time during a run. Thus it is sensitive to sudden bursts of the decay rate of the radio-active source which results in changes of the target generation time. Based on the model that PK involves sudden and short breakouts of action as demonstrated in the PKMB research (see Hasted 1982 and Betz 1978) one would expect that the variable ZTA shows strong correlations to the psychological variables.

H1: The test values ZTA shows stronger correlations to certain psychological variables than the test value T.

ri(ZTA) > ri(T)

H0: No psychological variable shows a stronger correlation to ZTA than to T.

8. The hypothesis ´hits or variance´ (16) was formulated on the basis of the experimenter´s subjective impression during the pilot experiment that subjects seem to ´produce´ either deviations in the hit rate or in the variance of the hit distribution but not in the two simultaneously. If this could be corroborated one could assume a model in which the ´PK-action´ could ´use´ two independent ´orthogonal´ channels: ´hits´ or ´variance´. Further, it could be assumed that the selection of this ´PK-channel´ is more or less arbitrary and that only the integral ´PK-action´ depends on psychological variables. The variable ZB (8) was designed to weighten only either deviations of V (2) or T (1) but not the two together. Of course one could doubt such an assumption, but by comparing the correlations of the psychological variables with T, V and ZB one could decide whether the selection of the channels T or V does in fact depend on psychological variables. It seems plausible to assume that such a possible integral ´PK-action´ depends primarily on the sheep-goat variable since ´goats´ might

(subconsciously) try to suppress not only deviations in T but also variance effects (increased V) which they could interpret as momentary ´successes´.

H1: The testvalue ZB yields stronger correlations to the sheep-goat variables than T or V.

rSG(ZB) > rSG(T) or
rSG(ZB) > rSG(V)

H0: The testvalue ZB does not yield stronger correlations to the sheep-goat variables than T or V.

9. The fluctuation hypothesis (17) is related to the hypothesis of target generation (15). The corresponding testvalues ZC (9) and ZD (10) describe different types of fluctuations which can occur in the RNG system. ZC is a measure which puts more weight on hits which do not coincide with a joint fluctuation of the source. This also holds for ZD. However, with this variable the criterion for a joint fluctuation is much weaker. Thus, if it is assumed that the influence hypothesis (9) is wrong a PK-effect could only occur in the display channel. In this case we would naturally expect that ZC shows stronger correlations than ZD. The two, however, should show stronger correlation than T because ZC and ZD have the same meaning as T except for the different weighting factors which emphasize fluctuations in the display channel. For this reason we would even expect for ZC and ZD the same psychological relationships as for T and therefore we can use the sum of the corresponding correlation coefficients for the operationalization of the hypothesis. It is obvious that this hypothesis does not make sense for the Schmidt type RNGs.

H1: The testvalues ZC and ZD show stronger relationships to the psychological variables than T:

SUM ri(ZC) > SUM ri(ZD) > SUM ri(T)

H0: The relationship of ZC, ZD, and T does not fulfil the given inequality.

10. The hypothesis of ´chaotic influence´ (18) is also based on experiences during the pilot study. It turned out that a special subject who had formerly been engaged in an alleged RSPK-case as the supposed focus person ´influenced´ the RNG system in an unpredictable

way. In comparison to control runs the experimental runs showed about twice times more significant deviations in test quantities, but also for the non-display channels. It seemed as if the subject was able to 'irritate' the whole RNG-system in various ways, because from run to run the variables which exhibited such deviations changed. The only consistent finding was that the total number of significant deviations was increased in comparison to the control runs. It turned out that the test quantity ZE (11), which is simply the sum of all five auto-correlation test quantities, is well suited to measure this effect. Of course it is difficult to formulate a precise hypothesis about such unexpected behaviour. However, similar to the variance hypothesis (13) we simply expect that ZE should at least for some psychological variables show higher correlations than T.

H1: ri(ZE) > ri(T)
H0: No psychological variable show a stronger correlation to ZE than to T.

As will be further discussed in paragraph 7, the validation of the hypotheses formulated here and in part II cannot be done in a completely formalistic way as is suggested by the formulation in disjunct sets of H1 and H0. The main reason for this lies in the circumstance that the different hypotheses are not independent and that the dependencies cannot be formulated in a quantitative way. Thus the rejection of one hypothesis may influence the interpretation of another one. One should also take into account that in a statistical experiment significant but meaningless random correlations might show up. Of course it is possible to estimate the number of such random correlations but only in the case the corresponding variables are totally independent or when their intercorrelations are known. Moreover it cannot be expected that PK yields strong effects. Thus the acceptance of a hypothesis should not exclusively depend on a given p-value. For instance, if it should turn out that the influence hypothesis is rejected because it failed by a fraction to reach the five percent level, whereas the other dependent hypotheses support the assumption of an influence, and if in addition the whole qualitative structure of these correlations fits with the influence hypothesis then one should really accept the hypothesis. On the other hand one has to be very careful with the application of such 'soft' criteria, especially if they are used post hoc. Hence it follows that the evaluation of a hypothesis should be based on both global and differential statistical tests.

3. EVALUATION METHODS AND PROCEDURES

To evaluate the hypotheses discussed in this paper and in Part II it is necessary to establish the evaluation methods in advance in order to avoid statistical selection artefacts (see Timm 1983b). However, it may well be that in the course of the evaluation it becomes apparent that certain methods cannot be applied due to a lack of required presuppositions (for instance with the application of parametric tests). In such a case it is necessary to ensure that the choice of the method which is then applied does not depend on results of the previous evaluations.

Timm (1983) has argued that in parapsychological research it is important to establish first whether there exists a significant overall result which can be interpreted as a 'psi'-effect, before differential methods are to be applied. Since this requirement seems rather restrictive - as will be discussed later - it was decided to use global and differential methods more or less as complementary approaches. The main objective in our approach was to keep the entire evaluation as transparent as possible and to use only methods which allow a clear interpretation of the results. This restriction is necessary due to the enormous amount of variables and aspects which could be investigated on the basis of our data. This means that further post-hoc studies are planned and that the evaluation methods reported here can be considered as a first approximation.

Formally the evaluation procedure involves 20 hypotheses concerning 31 independent and 18 dependent variables under four different experimental conditions. In table 1 these variables and the experimental conditions are summarized. The dependent variables were calculated for each of the 10 parallel random sequences ($i = 1,.....,10$) of each of the 8 runs ($j = 1,.....,8$) per subject with the exception that the variables K' and ZC can only be defined for the Markow sequences and K', ZC, ZD and ZE make only sense for all five parallel channels of each run. Obviously the abundant amount of this data requires a method for data reduction.

3.1 The data reduction method

The data reduction involves two aspects. The simpler one involves

TABLE 1
The variables and the experimental conditions

independent variables		dependent variables	
symbols	names	sym	names
PRIOR1-	confidence-score	T	number of hits
PRIOR7	before run 1,3,5,7	V	variance
POST1-	confidence-score	R	run score variance
POST7	after run 1,3,5,7	D	Kolmogorow-Smirnow
SGFAK1	personal exp. ESP	K	auto-correlation
SGFAK2	belief in own PK	(K′)	cross-correlation
SGFAK3	motivation to do PK	ZTA	generation time
IPCFAK1	externality	ZB	hits or variance
IPCFAK2	internality	(ZC)	joint fluctuations
IPCFAK3	fatalism conc. cars	ZD	time fluctuations
FPISK1	nervousness	ZE	chaotic influence
FPISK2	aggressiveness	ETS	positive novelty
FPISK3	depressivity	BES	duration of ETS
FPISK4	excitability	PIS	product of ETS.BES
FPISK5	sociability	ETG	negative novelty
FPISK6	calmness	BEG	duration of ETG
FPISK7	dominance	PIG	product of ETG.BEG
FPISK8	inhibition	DIM	prag. inf. display
FPISK9	openness		
FPISK10	extraversion	Markow RNG display on	
FPISK11	lability		
FPISK12	masculinity	Markow RNG display off	
EWLFAK1	non active mood		
EWLFAK2	depressed mood	Schmidt RNG display on	
EWLFAK3	high spirits		
EWLFAK4	nervous mood	Schmidt RNG display off	
EWLFAK5	frustration		
number j of runs per subj.: 8		experimental conditions	

the rejection of data or the combination of corresponding sets of data. Thus, as already mentioned, in part I only every second non-feedback run was stored and hence evaluated (except the variable T which was evaluated for all runs). Furthermore the two runs with the same experimental conditions (run 1 and 5, 3 and 7, 2 and 6, 4 and 8) were added together in such a way that the corresponding variables were summed up. As a consequence the psychological variables PRIOR3 to PRIOR7 and POST1 to POST7 had to be left out as predictor variables for the parapsychological aspects. Of the ten parallel random sequences only the data of the display channel were used except when data from the parallel channels were needed to calculate the variables K´, ZC, ZD and ZE. In the case of the non-feedback runs data were used from the same channel as used in the previous feedback run. Thus feedback and non-feedback runs were treated in the same way (but the sample size of the non-feedback runs was only half of the feedback runs).

Another more complicated method of data reduction is related to the aggregation hypothesis (8). This hypothesis was formulated in part II in relation to the psychological variables according to Brunswick´s lens model. The model also requires an aggregation of the dependent physical variables. However it is not clear how this should be done because, as stated before, we do not know the correlations between the physical variables (especially not under the PK hypothesis).

On the basis of the results of the pilotstudy, which indicated that the influence hypothesis has to be refuted, we decided to use the parallel random sequences which was not displayed as a frame of reference to solve the problem of aggregation of the physical variables. This should be regarded as a first approximation to the problem and not as a final solution. The underlying idea runs as follows: If the feedback hypothesis is correct and if the influence hypothesis is wrong the parallel random sequences can be regarded as control sequences. The intercorrelation pattern between its corresponding physical variables could then be regarded as the intercorrelation pattern under the null hypothesis. As a first approximation we could aggregate the corresponding physical variables of the display sequences in the same way by using the correlation coefficients of the ´control´ sequences as weightening factors (factor loadings). This model implies that a possible PK effect exhibits the same ´internal structure´ as pure randomness. This assumption might be highly questionable from the point of view of the classical (Rhinean) paradigm which states that PK is an additional process (or influence) to the random process which can be distinguished from chance behaviour. However, if the influence hypothesis has to be refuted it might well turn out that PK and chance events cannot be distinguished in principle (see Lucadou 1979, 1984, 1986).

In order to test the aggregation hypothesis both the original and the aggregated variables in the final evaluation were used.

Figure 1 shows schematically how the aggregation was done. As control sequences the parallel sequences of the third channel (i+2) from the display
channel (i) were used. Their corresponding variables were submitted to a factor analysis which led to a six factor solution. Table 2 shows the factors and their interpretations (names) and their corresponding marking variables for both the Markow- and Schmidt sequences. The factor loadings of the orthogonal rotated factor matrix were used afterwards to calculate the factors of the variables of the display sequences. For the non-feedback runs the factors of the variables were calculated with the same factor loadings (for details see v. Lucadou 1986). Thus again, feedback and non-feedback runs were treated in the same way. The advantage (if there is one) of this procedure is that the aggregation for both feedback and non-feedback is the same. However, as a result we have not obtained a real reduction of physical variables in our final evaluation but merely six additional physical variables. However, if the aggregation hypothesis (8) (Brunswick's lens model) and our special procedure of aggregation is correct we would expect increased correlations between the psychological variables and these factors.

3.2 Global evaluation techniques

As global evaluation technique Timm (1983b) proposes to apply a canonical analysis which should take into account all dependent and independent variables (including dummy variables for the partitioning of groups of each single variable) and also as variables all experimental conditions. The procedure of canonical analysis is an extension of the multiple regression technique. Linear combinations of one set of variables (independent variables) are correlated with linear combinations of a second set of variables (dependent variables) until an optimal correlation value is observed. Only if this global canonical analysis leads to a significant first canonical correlation

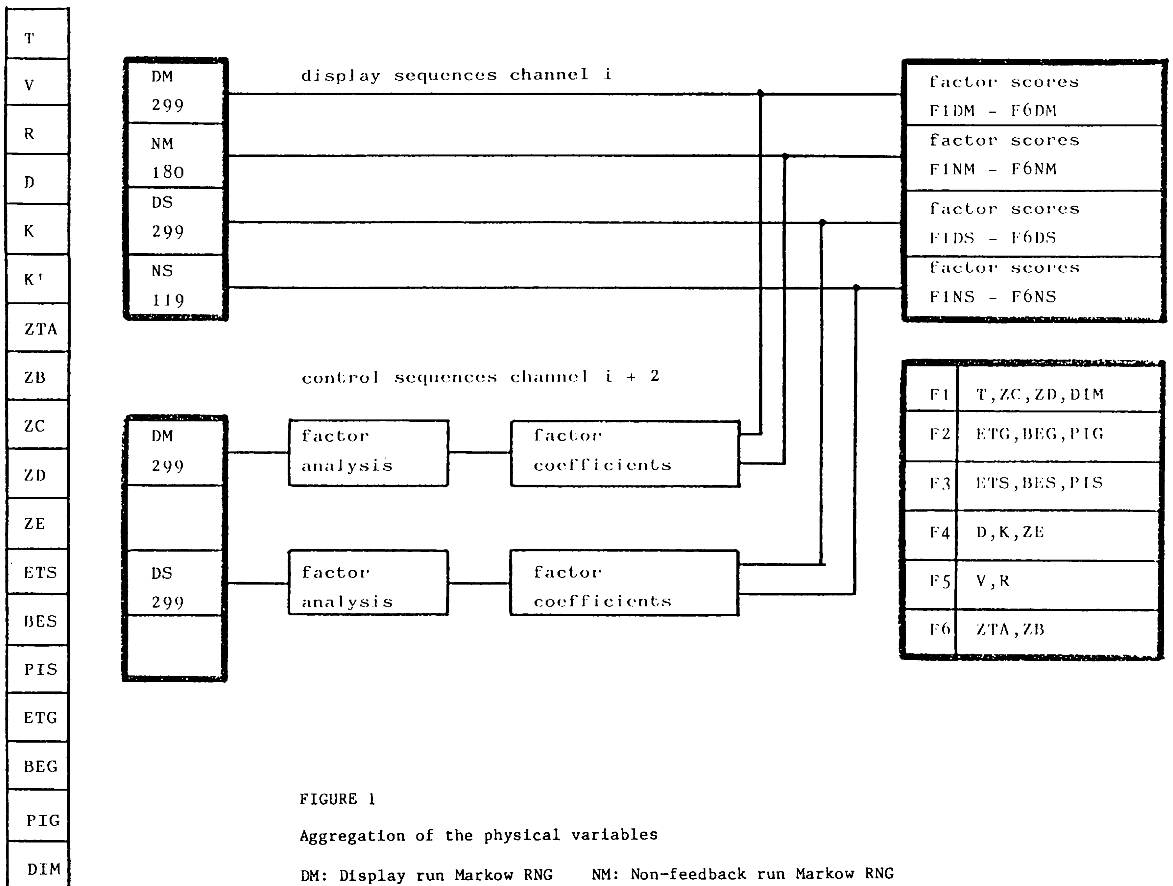

FIGURE 1

Aggregation of the physical variables

DM: Display run Markow RNG NM: Non-feedback run Markow RNG
DS: Display run Schmidt RNG NS: Non-feedback run Schmidt RNG

TABLE 2
The factors of the physical variables
and their associated variables

factor name	associated variable: Markow	associated variable: Schmidt
F1 hits	T,ZC,ZD, ETS,ETG, DIM	T,ZD, DIM
F2 negative fluctuation	ETG,BEG, PIG	ETG,BEG, PIG
F3 positive fluctuation	ETS,BES, PIS	ETS,BES, PIS
F4 distribution	D,K,ZE	D,K,ZE
F5 variance	V,R	V,R
F6 time - deviation	ZTA	ZTA,ZB

further inferential methods should be applied. According to Timm any differential evaluation of psi experiments whose data do not fulfil this requirement would be useless.

Apart from the technical problem of computer capacity when such a lot of variables is involved there still remains the more fundamental problem concerning the number of variables. In the case of a large number of variables Timm's method becomes very restrictive. If we assume for instance that only one pair out of hundred variables would yield a 'real' significant PK-correlation on the .01 significance level, this correlation would totally disappear in the 'noise' of the other variables. Consequently Timm requires that experimental designs with 'too many' variables should be avoided. Obviously this is not a clear criterion since especially in PK-research neither psychological nor physical relationships are sufficiently established to allow the a priori selection of 'relevant' variables. On the other hand the model of Brunswick's lens explicitly requires as much variables as possible in order to aggregate the 'relevant' ones. Thus Timm's requirement is rather fruitless in parapsychological experiments especially if an experiment is run the first time and is of an exploratory nature. It seems that the method can only fruitfully be applied to replication studies in which the 'relevant' variables could be designated beforehand. From this point of view one can argue that the study described here is only of heuristical value. However, by the combination of global and differential evaluation methods which where established beforehand it is intended to circumvent this fundamental methodological problem.

Moreover, even in the case that such a global canonical analysis would have been feasible and assuming that a significant canonical correlation was observed it would be rather difficult to validate the p-value. One could argue that the missing parametric assumptions could have caused an artificially high or significant canonical correlation. The canonical correlation is the maximized correlation which can be obtained from a suitable selected linear combination of independent and dependent variables respectively. For these reasons, only canonical analyses were calculated separately for the four experimental conditions using the 23 psychological and 17 physical variables.

Another global test which was used cannot settle the PK hypothesis directly but is linked to the feedback hypothesis (2). It is assumed that at least the aggregated physical variables F1 to F6 will show a global significant difference between the feedback and the non-feedback conditions for both Markow- and Schmidt type RNGs. However, since for the calculation of the factor values Fi Z-scores were used and because the distribution of the independent variables is well balanced one cannot expect a difference in the mean values but only in higher statistical moments.

Since the independent psychological variables are balanced, for instance, equal numbers of sheep and goats, any deviation of the total number of hits from the expected value must be due to biases of the RNGs.

3.3 Differential evaluation techniques

Since the distribution of many physical variables under the null hypothesis is not known only very simple non parametric methods were used for the evaluation of differential relationships between psychological and physical variables.

To cover both linear and non-linear relationships for all pairs of dependent and independent variables including the aggregated physical factors Fi Spearman's rank correlation coefficient and rank variance analysis (Kruskal-Wallis) were applied. For the rank correlation coefficients a significance level of 1% and for the rank variance analysis a significance level of 5% was chosen. For the analysis of variance a multiple comparison of the groups was carried out. The partitioning of the independent (psychological) variables into groups was the same as for the investigation of the psychological aspects as described in part II.

Several questions of differential nature could not directly be submitted to a statistical test procedure. Nevertheless quantitative and qualitative comparisons of corresponding variables can often be used to accept or reject a hypothesis or a model even if no p-value for the significance could be given. This is especially the case if the underlying model predicts a specific structure of the results. For instance, if it predicts that a certain physical variable should show the same relationship to a set of psychological variables as another physical variable. In such a case the 'structural' information which cannot be expressed in p-values becomes more important than the quantitative one. Very often such structural information is underestimated by the statistical methods. Thus Wittman (1983) emphasizes that such multivariate methods as factoranalysis, multiple regression analysis, canonical functions and discriminant functions should not be used 'blind'. This means that they should not be used without underlying theoretical model which allows a proper interpretation of the results. Furthermore many of these multivariate techniques use stepwise approaches, which implies that the methods optimize certain structures which can be found post hoc in the data. For this reason as Timm (1983) has pointed out it is highly questionable to use these methods as a strategy to prove the existence of an effect in parapsychology. As a consequence it was decided to make this problem transparent by only using the correlation matrices and to discuss all hypotheses on this basis without using further statistical multivariate methods.

4. BIAS OF THE RNGs

In parapsychological experiments the problem of a possible bias of an RNG is familiar and different experimental techniques have been used to rule out systematic errors. Such techniques are especially important in the case of proof-oriented research. In differential experimental designs a possible bias is automatically ruled out because only the difference of the outcome of a RNG under different (psychological) conditions is considered.

The reason for an RNG bias can be manifold. For instance, Bierman (1982) supposed that RNGs, which use flip-flops as circuit-elements like Schmidt-type RNGs do, could show a bias because certain commercially available types of flip-flops may have different error probabilities or switching sensitivity in relation to the two binary output states. For the Markowian RNGs a bias could be produced due to the finite decay time of the radio-active source or to a slow increase or - more likely - to a slow decrease of the sensitivity of the Geiger-Mueller tubes. In the case of our two types of RNGs the electronic circuit was the same for all five channels. Therefore we expect that any possible bias of the parallel RNGs will show the same tendency.

Since the present experiment is primarily a differential one, we only tested the RNGs in respect to possible bias with a sample size in the order of magnitude of a single experiment per subject. As described in more detail in (Lucadou 1986) no significant biases were observed. It is important to note that the length of the control runs to check whether the RNG is functioning properly should not differ too much from the length of the experimental runs, because it is possible that biases change after a certain period. For instance, it could well be that an RNG produces a bias due to thermal conditions, first in one direction but after a while in the opposite direction, or even one which oscillates between the two directions.

The usual techniques for eliminating the effect of such possible RNG biases, for instance by a systematic change of the target direction, where not used in this experiment because it would have eliminated the possibility to investigate the hypothesis concerning an possible physical PK-influence on the source (hypotheses 9, 10, 11, 12, 15, 17, 18). Especially the correlation technique would become rather insensitive if we would have to distinguish between a possible ´psi-tracer´ and an artificial auto-correlation produced by the systematic change of the target direction to balance out a possible bias. On the other hand such techniques to eliminate the effect of a possible bias does not make it vanish physically but simply conceals or even obscures it in relation to a given evaluation technique (for instance the binomial test).

Table 3 presents the total number of hits and the corresponding CR-value for the four experimental conditions. The sample size (number of subjects) is 299 for all conditions because the number of hits was stored for all runs. According to the symmetrical distributions of the independent psychological variables a large overall deviation from the theoretical expectation value is not expected. However, it appears from table 3 that especially the runs under the non-feedback condition show deviations which are partially significant. Since this result may indicate that the RNGs had a systematic bias a control run without a subject being present was carried out, consisting of 3182 normal runs. The result of these control runs are presented in table 4. In this case the sample size is much larger than that of the whole experiment. Since a systematic bias with a constant tendency is expected we could use the results of table 4 to calculate a bias correction for the overall scoring in the experiment. The corrected figures are presented in table 5. However, this bias correction can only be an approximation since the control run was carried out after the experiment was finished and the control run was carried out without interruption. This is of course a physical condition somewhat different from the real experiment.

The bias-corrected scores are given in table 5 separately for the different RNGs and feedback conditions. As a result it turns out, that in the non-feedback runs the significant deviations from the theoretical expectation value vanishes. Furthermore, it can be seen

TABLE 3
Total number of hits for the four experimental conditions

experimental condition	expected hits	observed hits	deviation	CR	p (two-tailed)
Markow (feedback)					
sum run 1, 3, 5, 7	179400	179494	+ 94	0.54	n.s.
Markow (non-feedback)					
sum run 2, 4, 6, 8	179400	179790	+ 390	2.26	0.024
Markow (together)					
sum run 1 - 8	358800	359284	+ 484	1.98	0.048
Schmidt (feedback)					
sum run 1, 3, 5, 7	179400	179543	+ 143	0.83	n.s.
Schmidt (non-feedback)					
sum run 2, 4, 6, 8	179400	179655	+ 255	1.48	n.s.
Schmidt (together)					
sum run 1 - 8	358800	359198	+ 398	1.63	n.s.
total experiment (F)					
sum run 1, 3, 5, 7	358800	359037	+ 237	0.97	n.s.
total experiment (NF)					
sum run 2, 4, 6, 8	358800	359452	+ 645	2.45	0.014
total experiment					
sum run 1 - 8	717490	718372	+ 882	2.55	0.011

from table 5 that, as expected, the bias of the RNGs shows the same tendencies for all five RNGs of the same type. It seems likely that the differences in the biases for each type of RNG only depends on product tolerances of the electronic and physical components but not on the setting of the counters.

The difference between the feedback and the non-feedback conditions of the corrected scoring rates still shows the same tendency as in the case of the uncorrected ones, but not to a statistically significant degree. Unfortunately it turned out, that the group frequencies for

TABLE 4
The measurement of the bias of the RNG

number of runs = 3182				
RNG	T	T-TO	CR	600.C
1 M	955991	1391	3.5	0.87
2 M	955571	971	2.4	0.61
3 M	955271	671	1.6	0.42
4 M	955418	818	2.1	0.51
5 M	956501	1591	4.0	1.00
1 S	956191	1901	4.8	1.19
2 S	957837	3237	8.1	2.03
3 S	955532	932	2.3	0.59
4 S	956622	2022	5.1	1.27
5 S	956415	1815	4.6	1.14

M = Markow RNG; S = Schmidt RNG; 600.C = Number of hits which has to be added in each run to correct for bias

the different RNGs are very unequal due to the unexpected behaviour of the pseudo random generator. Hence it is difficult to evaluate the accumulation hypothesis (12). There is no indication from the overall scoring that the difference between feedback and non-feedback condition depends on the setting of the counters. Moreover it was not possible to evaluate the differential aspects on the basis of the special RNG-number. Hence the accumulation hypothesis cannot be

TABLE 5
Bias corrected scores of numbers of hits of channels 1 to 5 of the Markow RNG and the Schmidt RNG

display-runs				non-display runs				
RNG	T-TO	n	CR	T-TO	n	CR	counter setting	count cycles
RNG1	-191.6	168	-2.09	10.4	168	0.11	152	1380
RNG2	89.8	177	0.95	65.8	177	0.70	192	1630
RNG3	-24.6	122	-0.32	88.4	122	1.13	112	1110
RNG4	70.7	106	0.97	63.7	106	0.87	32	340
RNG5	-43.5	25	-1.23	-31.5	25	-0.89	72	690
total	-99.2	598	-0.57	196.8	598	1.14	Markow RNGs	
RNG1	18.0	172	0.19	-130.4	172	-1.40	152	1380
RNG2	-117.2	151	-1.35	25.4	151	0.29	192	1630
RNG3	-42.5	115	-0.56	-4.4	115	-0.06	112	1110
RNG4	-10.6	52	-0.21	-68.3	52	-1.34	32	340
RNG5	-89.3	108	-1.21	48.1	108	0.66	72	690
total	-241.6	598	-1.40	-129.6	598	-0.75	Schmidt RNGs	
1-10	-340.8	1196	-1.39	-129.6	1196	0.28		

n = number of runs
T-TO = deviations from expected values

evaluated.

As already stated, the problem of RNG bias does not affect the differential aspects of the experiment. Moreover, from the last column of table 4 it appears that its effect on a single run is very small. Hence the differential aspects of the experiment were evaluated without bias correction.

5. THE RESULTS OF THE PILOT EXPERIMENT

Initially the aim of the pilotstudy was to find a specific process or tracer which could be used in the main experiment as a prominent PK-variable. Thus the pilot experiment was neither proof-oriented nor differential. In the pilot experiment the psychological conditions were described qualitatively and no fixed protocol was adopted. The main purpose of these experiments was to obtain a strong PK-effect. In some cases, however, formal sessions with a previously fixed number of feedback- and non-feedback runs were held. At the end the experimenter selected those runs which, to his judgement, could be considered as ´successful PK-runs´ together with a sufficient number of non-feedbackruns which served as controls. On the basis of these data it was tried to find a tracervariable which shows a clear difference between ´PK-runs´ and controls.

To illustrate how these experiments were done and which subjective criteria led to the conclusion of a ´successful PK-manifestation´ two typical cases are reported here:

a.) A student in arts came to the lab and reported about his personal ´psi-experiences´. He was very confident as regards having ´psychic powers´ but remarked that they were mainly related to musical events. A standard PK experiment with an optical display was run. After four ´unsuccessful´ runs the subject became rather frustrated and told the experimenter that he was sure to be able to influence the pitch of a tone ´psychokinetically´ but not ´such light bulbs´. To his surprise immediately another experiment was arranged with the same apparatus but this time with an acoustical display. However, the experimenter forgot to mention that in this version the feedback of the previous run normally presented during the intermission was switched off. After another four runs the subject stopped rather frustrated and complained about the unjust judgment during the pauses. He told the experimenter that he was convinced that he had indeed been very successful. It turned out that in fact the feedbackruns showed a CR-value of 3 whereas the non-feedbackruns showed no significant deviation. After the results were explained to him he tried another 7 runs, however, without further success.

b.) At a welcome party at the institute guests were invited to ´try their luck with the PK-machine´. Despite the relaxed atmosphere the ´subjects´ were not successful. Suddenly a young lady arrived who claimed to be able to ´influence the lights´. Since she was very confident and since many methodologically trained psychologist were present I decided to do a formal experiment. We decided to do 4 feedback-runs with 4 pauses. Unknown to the subject and to the other guests the pauses were used for a non-feedback run. Under the feedback condition she scored CR=3. However the non-feedback runs showed also a significant deviation of CR=2. After the experiment she disappeared, therefore further experiments could not be arranged.

As regards the psychological description of the ´successful´ subjects I can only report my personal impressions because no questionnaires were used in the pilot experiment. I found that persons who gave the impression of being extraverted, alert, intelligent, non-anxious, and rather confident about their abilities seemed in general to be more successful than those exhibiting the corresponding antagonistic characteristics. Another impression I had might also be worth to be reported: Those persons who appeared to be ´dogmatic´ believers and wanted to ´proof´ their ´abilities´ seemed to be the only ones which were really frustrated by the experiment; in fact they seemed to be less successful than the ´moderate sheep´.

For the quantitative evaluation of the pilot study the scoring rate could not be used as a criterion for a PK-tracer since this quantity was already used as selection criterion by the experimenter. The main methods which were used to find a tracer in the data, e.g. the ´correlation technique´ and the ´filter method´, are described in part I, chapters 4.3 and 5. The correlation technique primarily tests the influence hypothesis (9). If PK changes the decay rate of the radio-active source one would expect more similarity in the Markow chains of the five RNG-channels, and this can be detected by a decrease of the cross-correlation test quantity K´ (6). Table 6 presents the sum of these quantities for all four parallel channels of the four most successful ´PK-runs´ of the pilotstudy in comparison with subsequent non-feedback runs. It is obvious that a difference between the PK- and the control-runs is expected because due to the selection of a high hit rate the auto-correlation function of the Markow sequence must have been changed. If the decay rate of the source was affected one would expect a decrease but instead an increase of these values is observed. Therefore from the data of the pilot experiment we can reject the influence hypothesis (9). A visual inspection of the plotted correlation functions of some very successful PK-runs confirmed these findings.

TABLE 6
Sums of the cross-correlation test quantities K´ for ´PK-runs´ of the pilot study in comparison with control runs

RNG channel number	PK-runs	control runs
1	1920.28	1828.98
2	2075.65	1873.35
3	2353.22	1799.11
4	1920.59	1766.23
sum	8269.74	7267.67

To apply the filter method the data points of the ´PK-runs´ were plotted into the diagram (scattergram) (figure 7, part I) which shows the dependency between the number of hits (T) and the auto-correlation test values (K). It turned out that all points fit the curve which suggests pure random fluctuations. This must be seen as an indication that even the most successful ´PK-runs´ cannot be distinguished from normal chance fluctuations. However, without further research nothing can be said about the accuracy of this method.

In a similar way other combinations of variables were used to produce scattergrams which might yield different configurations for the ´PK-runs´ and the corresponding control-runs. However, no clear indication for such a special characteristics could be found for the two types of RNGs. As a consequence we have to reject the tracer hypothesis (10) for the pilot experiment.

6. THE RESULTS OF THE MAIN EXPERIMENT

Due to the fact that from the pilot study no tracer could be found many of the physical variables which were also used in the main experiment became obsolete since they were primarily introduced as possible candidates of a tracer. Not being a tracer, functioning as a measure for the strength of the PK-force or energy, most variables loose their meaning in the context of this experiment. For instance, it is not clear how any psychological variable could correlate with the variable D (which indicates deviations from the distribution under H0) because the instruction for the subjects did not include this variable. A deviation in D cannot necessarily be interpreted as a ´PK-success´ even if it is highly significant! The only physical variables which are clearly related to the instruction are those who belong to factor F1 (see table 2) which is called the ´hit-factor´. Nevertheless all variables were retained in the evaluation of the main experiment for heuristical reason although this might weaken the global evaluation techniques if one applies the criteria of Timm (see paragraph 3.2).

6.1 Qualitative Results

Qualitative results are results of analyses for which no p-value can be given but which confirm or reject a hypothesis on a structural basis. Here it mainly involves the influence and the tracer hypothesis. In this study the same results were observed as in the pilot experiment. Figure 2 presents as an example a comparison of the auto-correlation function of a feedback-run with the four cross-correlation functions of the parallel monitored channels. The selected run was highly significant as regards the number of hits (T = 322, CR = 3.1). Since for the other significant runs there is also no indication of any similarity between the auto- and the cross-correlation function, the influence hypothesis (9) has to be rejected. This result can also be seen from the data of table 7 which shows that the mean value of the cross-correlation test values (K´) for the six most significant feedback runs of the main experiment are significantly larger than the mean value of all runs. This is the same feature as was observed in the pilot experiment and proves that no common influence on the five Geiger-Mueller tubes has taken place.

TABLE 7
Cross-correlation test quantities K′ for the most significant feedback runs of the main experiment

mean value of K′ n = 598 (all runs)	STD-error	mean value of K′ n = 6 (extreme runs)
707.41	7.77	835.27

As regards the tracer hypothesis (10), figure 3 presents the corresponding scattergrams for both Schmidtian and Markowian feedback runs. Since all data points lie on the curve of pure chance fluctuations we have also to reject the tracer hypothesis.

6.2 Results of the global evaluation techniques

Since the evaluation of the overall hit rate did not show significant differences between feedback and non-feedback conditions (as discussed in paragraph 4), we should at least find a global effect in an aggregated variable if we want to adhere to Timm′s criterion. From the previous discussion it is clear that primarily an effect is expected for the factor F1, because this is the only factor which can be meaningfully interpreted in view of the instruction given to the subjects. We do not expect a deviation from the mean value because the distributions of the psychological variables are well balanced, thus ′psi-hitters′ and ′psi-missers′ would cancel out each others effects. Therefore we expect a difference in the variance (or higher statistical moments) of factor F1 between feedback and non-feedback condition. Table 8 presents the variance, the skewness and the excess for all six aggregated factors for both RNGs in comparison between feedback- and non-feedback condition. It turns out that only the factors F1, F2 and F6 of the Markow runs yield significant differences in variance. However for F6 the requirements for the F-test are no longer fulfilled because skewness and/or excess yield high values.

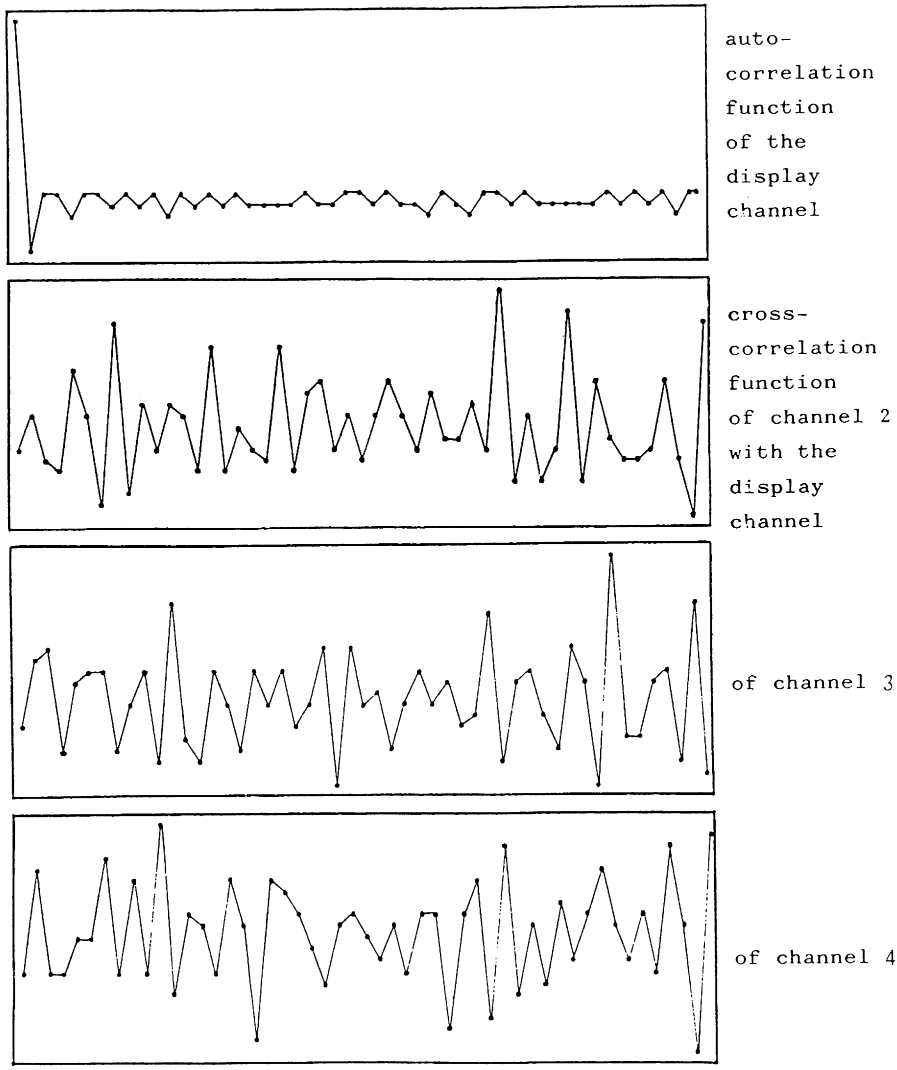

FIGURE 2

Correlation functions of a single run

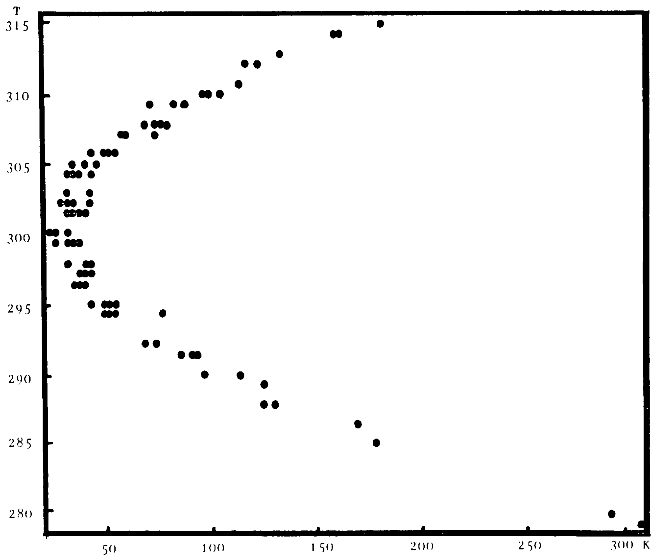

FIGURE 3a SCATTERGRAM (T versus K) for feedback runs (Markow RNG, channel 1)

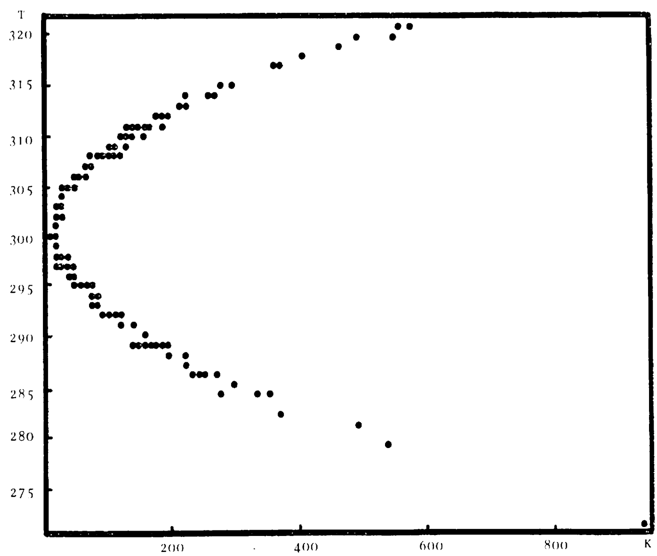

FIGURE 3b SCATTERGRAM (T versus K) for feedback runs (Schmidt RNG, channel 1)

These results support primarily the generator hypothesis (11) which predicts a difference between the two types of RNGs.

TABLE 8
Differences between feedback and non-feedback condition of the physical factors F1 - F6 of table 2

Markowian run (n1 = 299, n2 = 180):								
var	variance display	variance non-disp.	F	p	skewness display	skewness non-disp	excess display	excess non-disp.
F1M	2.145	1.596	1.344	.025	-.327	.070	.215	-.068
F2M	2.003	1.596	1.255	.050	.464	.280	-.081	.162
F3M	2.027	2.103			.585	.380	.718	-.117
F4M	1.828	1.959			.799	.949	1.237	.962
F5M	2.003	2.040			.502	.398	.508	-.324
F6M	2.757	2.007	1.374	.025	7.990	.289	95.945	-.313
Schmidtian runs (n1 = 299, n2 = 119):								
var	variance display	variance non-disp.	F	p	skewnes display	skewness non-disp	excess display	excess non-disp.
F1S	1.902	2.204			-.029	.037	-.221	-.219
F2S	2.050	2.027			.927	.639	.944	.310
F3S	2.075	2.173			.689	.252	.686	-.420
F4S	1.939	1.948			1.298	1.501	2.123	3.509
F5S	1.911	2.052			.121	.519	-.325	.472
F6S	2.010	2.105			2.190	.513	20.140	4.646

This difference between the Schmidtian and the Markowian RNG can also be seen in the canonical analysis which was performed for the four experimental conditions separately, with psychological variables as independent and physical variables as dependent ones. The procedure selects linear combinations of the dependent respectively independent variables in such a way that the correlation between these two linear combinations becomes maximal. The significance of this correlation can be used as a criterion according to Timm's approach. However, it turned out that only under the condition of Markowian feedback runs a tendency for a canonical correlation r could be found (r = .52, Wilk's lambda = .2, p = .1). It seems plausible that this poor result is mainly caused by the high number of 'irrelevant' variables. This assumption will be supported by the results of the differential analysis.

Especially in the case of the Schmidtian RNG it is not possible to conclude that a PK-effect is found in the data on the basis of the global evaluation technique. However, since Timm's criterion appears to be very restrictive it seems wise to base the final decision as regards the PK hypothesis also on the outcome of the differential tests.

6.3 Results of the differential analyses

The differential results are given by the four correlation matrices (tables 9 to 12) showing the relationship between psychological and physical variables, including the aggregate factors for all four experimental conditions. Each line of the matrices represents a psychological variable (factor) and its correlation to the physical variables. As discussed in paragraph 3.3 the two methods to find such relationships, namely rank-correlation coefficients and rank variance analysis, can be considered as complementary techniques. One emphasizing linear and bidirectional, the other also non-linear and unidirectional relationships. In the correlation matrices stars indicate significant variance analysis (one star: p = .05 and two stars: p = .01) and numbers indicate the correlation coefficients significant at p = .01. The sequence of the variables corresponds to the factors. It has to be taken into account that those (physical) variables which belong to one factor may depend among each other. However, this only holds under the null hypothesis. For the PK hypothesis there might be quite distinct relationships which only concern special pairs of variables.

In general the tables can be interpreted in the following way: If a line shows up many (and strong) correlations it means that the

psycho-logical variables	experimental condition Markow non-display physical variables F1	T	ZC	ZD	DIM	F2	ETG	BEG	PIG	F3	ETS	BES	PIS	F4	D	K	ZE	F5	V	R	F6	ZTA	ZB	H1/T	TNDZ
PRIORI																	*							?	?
SGFAK1														* -.193	** -.248	** -.201								?	?
SGFAK2																								?	?
SGFAK3														*	*									?	?
IPCFAK1																								?	?
IPCFAK2																								?	?
IPCFAK3																								?	?
FPISK1														**	*	*	*							?	?
FPISK2		* -.144																			**	**		?	?
FPISK3																								?	?
FPISK4																	* .197							?	?
FPISK5																								?	?
FPISK6							*																	?	?
FPISK7						*																		?	?
FPISK8	*			*	* -.184																			?	?
FPISK9																								?	?
FPISK10									*															?	?
FPISK11														*			** .202							?	?
FPISK12																								?	?
EWLFAK1														*	*									?	?
EWLFAK2																								?	?
EWLFAK3														* -.241	-.184	-.216								?	?
EWLFAK4					*																			?	?
EWLFAK5		*									*				*	*						*		?	?
number n	180	299	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180		

TABLE 9 Correlation matrix for non-feedback Markow runs

psycho- logical variables	experimental condition Markow display physical variables F1	T	ZC	ZD	DIM	F2	ETG	BEG	PIG	F3	ETS	BES	PIS	F4	D	K	ZE	F5	V	R	F6	ZIA	ZB	H1/T	INDZ
PRIORI					*																			+	+
SGFAK1										*.														+	?
SGFAK2																					**.	*		+	?
SGFAK3																								+	?
IPCFAK1					* .144						-.137													+	-
IPCFAK2							.			.				.										-	?
IPCFAK3														.		*								+	?
FPISK1	* -.169	** -.175	** -.174	* -.154	* .169					*	**	**	*											-	-
FPISK2																								?	?
FPISK3	* -.185	* -.187	* -.183	* -.170	** .203						* -.160													-	-
FPISK4	*.	**	**	*	**	*	**		*							*			*	*				-	-
FPISK5					*																			+	+
FPISK6		*	*		*						*		*											+	+
FPISK7										.		*												?	?
FPISK8	.	* -.141	*		** .174	.					* -.149			*				*						-	-
FPISK9																		.						?	?
FPISK10																								+	?
FPISK11											*		*											?	?
FPISK12	** .204	** .200	** .199	** .195	** -.224						** .197										**	** .136		+	+
EWLFAK1					*	*		*												*				-	-
EWLFAK2					**		*		*		*													-	-
EWLFAK3																								+	?
EWLFAK4						*			*													* -.139		-	?
EWLFAK5	.	.138	.142	* .155	.140						.163							** .204	** .214	** .198				+	+
number n	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299		

TABLE 10 Correlation matrix for feedback Markow runs

psycho- logical variables	experimental condition Schmidt non-display physical variables F1	T	ZC	ZD	DIM	F2	ETG	BEG	PIG	F3	ETS	BES	PIS	F4	D	K	ZE	F5	V	R	F6	ZTA	ZB	H1/T	TNDZ
PRIORI			--																			.212			
SGFAK1			--															*	*	*				?	?
SGFAK2			--																					?	?
SGFAK3			--																					?	?
IPCFAK1			--			*	*																	?	?
IPCFAK2			--																					?	?
IPCFAK3			--															*	*					?	?
FPISK1			--											**	*	*					.212	* .221		?	?
FPISK2			--			*				*	*	*	*											?	?
FPISK3			--																					?	?
FPISK4			--												* .269									?	?
FPISK5			--			*	*	*																?	?
FPISK6			--																					?	?
FPISK7			--			*																		?	?
FPISK8			--																					?	?
FPISK9			--												.213									?	?
FPISK10			--																					?	?
FPISK11			--																		*			?	?
FPISK12			--																					?	?
EWLFAK1			--																					?	?
EWLFAK2			--															**	**	**	*			?	?
EWLFAK3			--																			-.250		?	?
EWLFAK4			--							*			*											?	?
EWLFAK5			--												*						*	*		?	?
number n	119	299	--	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119	119		

TABLE 11 Correlation matrix for non-feedback Schmidt runs

psycho-	experimental condition Schmidt display																								
logical	physical variables																								
variables	F1	T	ZC	ZD	DIM	F2	ETG	BEG	PIG	F3	ETS	BES	PIS	F4	D	K	ZE	F5	V	R	F6	ZIA	ZB	HI/TI	INDZ
PRIORI	** -.180	** -.164	--	** -.164	** .179		.140	.153																+	+
SGFAK1			--											** -.200	** -.244									+	?
SGFAK2			--																					+	?
SGFAK3			--												*									+	?
IPCFAK1	*	*	--																					+	-
IPCFAK2			--																					-	?
IPCFAK3			--											**	**	*	**							+	?
FPISK1	.		--															.						-	?
FPISK2			--																					?	?
FPISK3			--																					?	?
FPISK4			--																		.			?	?
FPISK5		*	--	*																				+	?
FPISK6			--																					+	?
FPISK7			--																					?	?
FPISK8,			--																					-	?
FPISK9			--															.	*	*				?	?
FPISK10			--																			*		+	?
FPISK11			--														*							-	?
FPISK12			--							*	*	**	*					*	**	*				+	?
EWLFAK1			--																					-	?
EWLFAK2			--									*												-	?
EWLFAK3			--											* .160	*									+	?
EWLFAK4			--									-.139							*				*	-	?
EWLFAK5			--																		.			+	?
number n	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299	299		

TABLE 12 Correlation matrix for feedback Schmidt runs

corresponding psychological variable is relevant for the PK-process. If many (and strong) correlations are found in a row this implies that the physical variable involved is sensitive to PK.

A rough estimation indicates that for both types of RNG the number of significant correlations under the non-feedback condition does not exceed the number of correlations which would be expected by chance at the given significance levels (table 9 and 11). This may be interpreted as an indication that non-feedback PK-effects did not occur. Only the clustering of a few correlations around factor 4 may be of some interest.

Under the feedback conditions, however, the number of correlations is increased, especially for the Markowian RNG and for factor F1. If only the significant relationships are considered at the .01 level the difference in the number of correlations between non-feedback and feedback condition is 3/11 for the Schmidt- and 6/22 for the Markow-runs. This result supports both the feedback- (2) and the generator hypothesis (11).

A direct support of the relevance of feedback and also a support of the model of pragmatic information (display hypothesis (19) can be found in the fact that indeed the variable measuring the pragmatic information of the feedback the variable DIM, showed not only the highest number of significant correlations to the psychological variables but also consistently the strongest correlation coefficients. This can be seen in table 13 where the correlation coefficients between a selected number of psychological variables and all variables which belong to factor 1 (hits) are compared. Only psychological variables were selected which show at least one significant correlation to one of these variables.

For the physical variables which allow a meaningful interpretation (factor 1) the tendency of most of the correlations (at least for the Markow runs) is in agreement with the expectations of our psychological hypotheses. This is indicated in the last two rows of the tables. The row labeled H1/T shows the expected tendency of the correlation ('+' for positive correlations and '-' for negative ones, and '?' if no hypothesis was formulated) and the row TNDZ shows the measured direction of the correlation (if no significant correlation was found this is indicated by '?').

For the discussion of the psychological hypotheses (which were formulated in part II) graphs of the multiple variance analysis are presented in figure 4. The abscissa shows the group division of the psychological variable in question and the ordinate the rank mean values of the variable DIM which has the closest correspondence to the instruction given to the subjects (as discussed in part I). The expected (hypothesized) tendency is indicated by a punctuated diagonal line. One should take into account that the variable DIM implies that low values correspond to high 'PK-success' and vice-versa. However, due to our differential method (and the problem of bias) we cannot apply the distinction between 'psi-missing' and 'psi-hitting'. Thus 'PK-success' can only be validated in relation to the other groups of a variable.

7. DISCUSSION

At first glance it seems necessary to start with settling the PK hypothesis (1) before other hypotheses can be discussed. However, this implies that already something is known about the nature of PK. This means that, on one hand, we cannot isolate the experiment from previous ones but, on the other hand, should not regard the experimental results as final answers. If for instance only one of the differential hypothesis could be confirmed due to a single significant correlation in our correlation matrices and if this single correlation is in agreement with the results of many previous experiments we should regard it as a confirmation and accept the PK hypothesis even if the overall results does not show a significant effect. Therefore we cannot settle the PK without referring to previous findings.

We will discuss first the five main hypotheses before we turn to the more detailed ones:

7.1 PK hypothesis (1)

Although there is at least one significant overall result for the Markow runs (in factor 1) one should not evaluate the PK hypothesis on this basis alone. However, the obvious difference between the feedback and the non-feedback conditions in combination with the fact that most of the psychological correlations are in agreement with the results of a previous independent experiment (Weis 1972) (see below) allows to

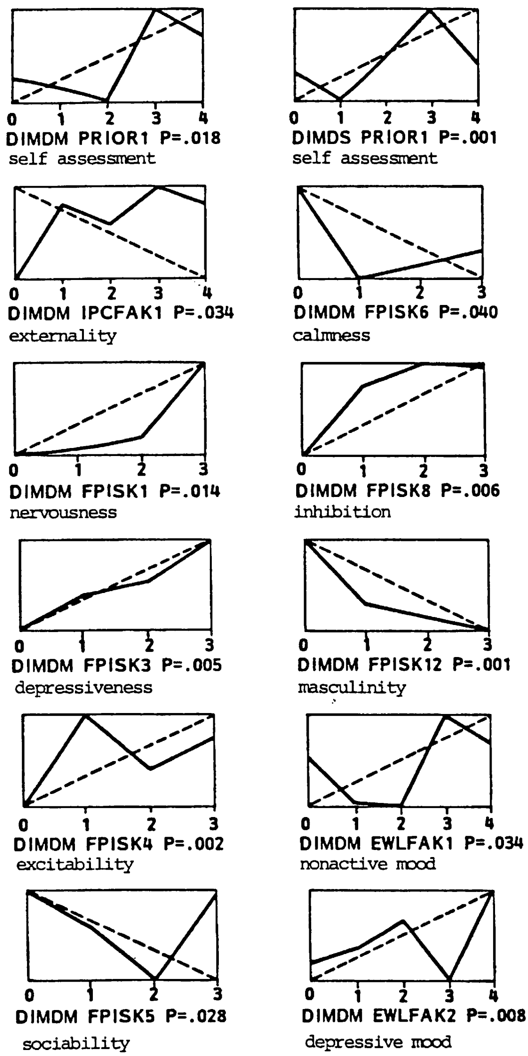

FIGURE 4

Graphs of the multiple analysis of variance. The expected tendency is indicated by the dashed line. Variables DIM and PRIOR1 are reversely poled. DM: Display Markow RNG; DS: Display Schmidt RNG accept the PK hypothesis at least for the Markow condition. For the Schmidt RNG it is much more difficult to evaluate the PK hypothesis because the relationships between psychological and physical variables are much weaker. In general they show the same difference between feedback and non-feedback condition. Moreover, some of the variables even show the same tendency although it has to be emphasized that the whole structure of the correlation matrix for the Schmidt RNG differs from the Markowian one. Since the subjects could not distinguish between the Schmidt and the Markow condition it seems implausible to reject the PK hypothesis for the Schmidt condition. However, in this case one should not attach too much value to the discovered relationships.

7.2 Feedback hypothesis (2)

For the Markow runs a significant difference between feedback and non-feedback condition is observed for the factors 1, 2 and 6. Of these factor 1 is the most important one. This result is supported by the consistent finding that the variable DIM shows the most significant difference between these two conditions. Since DIM quantifies the meaning of what the subjects actually observed on the display in relation to their instruction this is a direct confirmation of the feedback hypothesis. Again this also holds for the Schmidt runs but, however, to a much smaller extent. On the other hand, the clustering of correlations in factor 4 and 6 under non-feedback condition should not be overlooked. If this indicates a non-feedback PK-correlation this could be interpreted as a kind of ´linger´- or ´relaxation´-effect which may have occurred during the pauses when the subjects were not aware of the ongoing run. Such ´linger´-effects have been described in several RSPK-cases (see Roll 1977, Moser 1966, Bender 1977).

7.3 Generator hypothesis (11)

For the evaluation of the generator hypothesis one has to take into account that several physical variables cannot be directly compared for the Schmidt and the Markow RNG. For instance ZC cannot even be defined for the Schmidt RNGs. But for the variables which are comparable (T, V, R and DIM) it was observed that the differences between the two types of RNGs express themselves not only in the number of correlations or magnitude of the 'PK'-effect but also in a different structure of the relationships between physical and psychological variables. Therefore the generator hypothesis has to be rejected.

This is of great interest from a theoretical point of view. In Schmidt's model it is assumed that a PK-effect on a quantum mechanical RNG only depends on display and psychological conditions but not on the 'complexity' of the RNG. In Walker's model and the model of pragmatic information (MPI), however, the size of the effect may well depend on the physical process, namely the divergency of the process. The divergency of a quantum mechanical state can be described as the cardinality of the subspace to which the state vector collapses. The selection of the subspace may depend on physical as well as on psychological variables (see Lucadou 1987). Without further research it is a matter of speculation why the Markow RNG turned out to be more 'PK-sensitive' in this experiment. However, a restriction has to be made: Since the variance of the Markow chains is three times smaller than that of a normal random sequence the light column of the display showed somewhat weaker fluctuations for the Markow RNGs. Normally this could not be detected by the subjects, but we cannot rule out with certainty that this did not produce a different psychological condition, for instance a higher 'frustration' of the subjects during the Markow runs.

7.4 Influence hypothesis (9)

The influence (9) and the tracer hypothesis (10) has to be rejected. This is in agreement with the results of the pilot experiment. However, since in both cases the same methods were used, an independent support by other evaluation methods would be helpful, because this finding is of utmost theoretical importance. Fortunately, after the termination of this study a study of May et al. (1985) appeared which independently showed the same result based on a post hoc analysis of a great number of published PK experiments. The authors used a special statistical method to test the action of different models of 'PK-influence' on RNGs. They came to the conclusion that no indication for any physical influence could be found although the overall statistical significance for PK was overwhelming (see Radin and May, 1986). This result led the authors to the so-called ´intuitive data selection (IDS)´ model which assumes that PK-subjects ´select´ those random fluctuations from an otherwise normally distributed random source which are favorable for their task. They speculated that this ´selection´ might be similar to a kind of GESP. In a recent article Walker (1987) showed that the results could also be explained from the quantum mechanical observer model. In another recent article (see Lucadou 1987) it was argued that any ´real physical PK-influence´ would contradict both Walker´s model and the MPI model in which PK is assumed to be a non-local correlation in the sense of the EPR-correlation.

Based on our data we can rule out the GESP-assumption of the IDS model. If the subjects could precognitively foresee fluctuations in a ´prestabilized´ sequence of random events which they then select for their own purpose by pushing the start-button in the right moment, it can be expected that random sequences which show large fluctuations at the display would be a better target that those with small fluctuations. Hence, the subjects should be more successful with the Schmidt runs than with the Markow ones. As pointed out above, we found the opposite.

The rejection of the influence and tracer hypotheses has also some interesting methodological aspects. Since we could not find any signal or physical influence on our RNGs, we can also rule out the possibility that some unknown ´normal´ physical influence (such as temperature, electrical discharges, vibrations etc) of our subjects have surmounted the shielding of the apparatus and thus caused artificial correlations. As a conclusion we can formulate the seemingly paradoxical statement that the absence of a ´PK-signal´ seems to establish the existence of ´PK´.

7.5 Aggregation (8) and accumulation hypothesis (12)

In contrast to the ´physical´ hypotheses discussed above the aggregation and the accumulation hypotheses can be conceived as ´methodological´ ones. They have in common that they both start with the assumption that weak effects can be statistically accumulated to a significant effect. This assumption involves that random or chance effects cancel out when they are added but that ´genuine´ effects are really added. This model does not include a classical causal model and, hence, is also valid for quantum effects. However we cannot a-priori assume that this is also valid for psi-effects (see Lucadou 1984).

The aggregation hypothesis emphasizes the amplification of the PK-correlations according to Brunswick's lens model, whereas the accumulation hypothesis relates to the amplification of the 'PK-effect' on the random process.

The latter would be especially powerful in the case of a real PK-influence (hypotheses 9, 10), but also under the assumption that PK is 'only' a non-local correlation (see Lucadou 1987) one could imagine a different PK-effect on RNGs with different settings of the counter (see part I). In principle this is another type of generator hypothesis. Unfortunately it turned out that the group frequencies for the different RNG channels did not allow a straight forward test of this hypothesis. Furthermore, a comparison of the overall hit rates between feedback runs and non-feedback runs is obscured by the bias of the RNGs. Therefore this hypothesis cannot be evaluated.

The aggregation hypothesis (8) involves two aspects. One concerns the psychological variables, the other the physical variables. The effect of the aggregation of the psychological variables is not investigated in detail since only already aggregated psychological variables (factors and scales) were used and not single items. Nevertheless, the fact that FPISK12 (masculinity) showed the strongest PK-correlations can be interpreted as a support of this hypothesis since this scale is composed of the subscales FPISK1 (nervousness) and FPISK8 (inhibition).

As regards the physical variables the PK-correlation of single variables can be compared with the correlations of the factors. But since the factors were not aggregated on the basis of the 'PK-runs' but merely on the basis of the control runs (see paragraph 3.1) the method might be too conservative (unfair) in relation to Brunswick's lens model. In general these factors do not show higher correlations than the single variables. Nevertheless, they show a stronger difference between the feedback and the non-feedback conditions than any single physical variable (see table 8).

PK correlations between physical and psychological variables can be based either on a single run of the two runs with the same RNG condition in a session or on the sum of these two runs. The aggregation hypothesis predicts that the correlations based on the sum of the runs are stronger than those based on single runs. Unfortunately due to the unequal group frequencies of the different channels the analysis using single runs could not be carried out in a comparable way to those based on sum of runs. Nevertheless, in so far evaluations based on single runs could be meaningfully interpreted, the results show similar correlations. However these correlations are distributed irregularly over the runs of a session (see Kunzmann 1984). Only when equivalent runs of a session are put together a more clear picture emerges. Therefore we are inclined to accept the aggregation hypothesis in spite of the low numerical values of the correlation coefficient of the physical factors.

7.6 Discussion of the more specific hypotheses

The remaining physical hypotheses (13, 14, 15, 16, 17, 18, 19, 20) and the psychological hypotheses (3, 4, 5, 6, 7 formulated in part II) are much more specific and have less theoretical relevance compared to the general hypotheses discusses above. However, the acceptance of the PK hypothesis also depends on the acceptance of these more specific hypotheses.

Since the remaining physical hypotheses are of less general interest we will discuss them only briefly. A more detailed discussion can be found in Lucadou 1986.

The variance hypothesis (13) is based on the assumption that certain psychological variables may be good predictors for those subjects which do not produce a stable PK-effect but whose ´PK-performance´ fluctuates in time. In such a case one would expect an increase in the variance of the random sequence rather than in the number of hits. Indeed for EWLFAK5 (frustration) such a correlation to the ´variance variables´ F5, V, R, of the Markow RNG could be found which was stronger than the corresponding ´hit variables F1´. The correlation indicates that frustrated subjects produce larger variances than non-frustrated ones. Psychologically this seems plausible. For the Schmidt RNGs a non linear relationship of FPISK12 (masculinity) with V was found which, however, is difficult to interpret. Therefore the variance hypothesis has to be accepted in spite of the fact that the variance variables turned out to be not very sensitive to PK.

The distribution hypothesis (14) concerning the variables F4 and D has a similar rationale as the variance hypothesis, but it is much less specific. For this variables several correlations were observed under non-feedback conditions which are hard to interpret (see above). For the Schmidt runs relationships to SGFAK1 (personal belief) and IPCFAC3 (fatalism) can be found which also cannot be easily interpreted because the meaning of D is so unspecific. Formally, however, the hypothesis has to be accepted.

The variable ZTA (generation time) concerning the hypothesis of target generation (15) shows similarly irritating features as D, but the hypothesis can be rejected since the only correlation (in Markow runs) is smaller than the corresponding correlation to T.

The hypothesis ´hit or variance´ (16) can clearly be rejected. This means that in PK-runs T and V shows no ´exclusive or´ relationship. For the Markow runs this is explicit insofar as EWLFAK5 (frustration) shows positive correlations as well to the hit variables as to the variance variables.

The fluctuation hypothesis (17) should be evaluated by comparing the variables T, ZC and ZD in table 13. This table indeed shows that the correlations of the psychological variables to ZC are stronger than to ZD. At least it seems interesting that the difference between ZC and ZD is in agreement with our predictions. However, they are not stronger than the correlations to T as was expected. The fluctuation hypothesis is of more general interest since it presents a ´positive´ alternative to the tracer hypothesis. If it should be confirmed that the PK-effect is no ´signal superimposed´ on a random sequence then alternatively one has to assume that fluctuations of the system can be either ´selected´ as IDS (model of intuitive data selection) supposes or actively ´reinforced´ as the MPI (model of pragmatic information) predicts (for further discussion see Lucadou 1987). In any case one would assume that (the region of the state-space which can be adopted by) these fluctuations depend on the physical process, respectively its divergency, as is also assumed in Walker´s model.

It is also difficult to evaluate the hypothesis of ´chaotic influence´ (18). In contrast to our expectations based on the data of the pilotstudy it turned out that the variable ZE was not very sensitive and also not very specific. Its few correlations occur in the non-feedback situation and are therefore difficult to interpret. One might consider this as qualitative support of the idea that RSPK-focus persons ´produce´ quite different effects as ´normal´

TABLE 13
Comparison of the correlation coefficients of the variables of factor F1 of table 2 with psychological variables

Markowian feedback-runs				
psych. var.	DIM	T	ZC	ZD
IPCFAK1	.1435	-.1158	-.1044	-.0856
FPISK1	.1692	-.1748	-.1739	-.1537
FPISK3	.2027	-.1869	-.1830	-.1702
FPISK8	.1742	-.1405	-.1319	-.1061
FPISK12	-.2237	.1998	.1994	.1952
EWLFAK5	-.1395	.1384	.1423	.1545
Sum /r/	1.0528	.9562	.9349	.8653

Schmidtian feedback-runs			
psych. var.	DIM	T	ZC
PRIOR1	.1789	-.1640	-.1635

subjects but, of course, it is very speculative to draw such conclusions from a few correlations.

It is obvious that the variable DIM does play a predominant role, not only in view of the MPI but also considering the design of the experiment. Table 13 and the correlation matrices clearly indicate that the hypothesis of pragmatic information of the display (19) has to be accepted even though we cannot provide a p-value for the differences between the variable DIM and T. It is important to emphasize that DIM primarily operationalizes what the subject actually observes on the display.

Variable ETS (positive novelty) shows a rather similar behaviour as

DIM which seems plausible because it quantifies ´positive´ fluctuations which the subjects can observe on the display. It is interesting to find that ETG (negative novelty) does not show an antagonistic behaviour as one would expect at first glance. This may indicate that fluctuations in the ´negative´ direction have a different meaning for the subjects.

Since the experimenter does not observe the display in the session his only criterion for judging an experiment to be successful is the variable T. If we assume that the observed PK-effect is ´only´ an experimenter effect in the sense that the experimenter would consciously or subconsciously ´attribute´ a certain ´PK-success´ to certain subjects (for instance, according to his or her sympathy) then we would expect that the variable T yields the strongest correlations. But the experimenter did not take notice of the T values which were printed out during the experiment as a back-up for the data which were kept stored on disk. It can also be ruled out that the experimenter did this ´attribution´ during the course of the evaluation, because between the experiment and the evaluation period there was a considerable time interval of more than a year, hence the experimenter could not remember single experimental sessions. Furthermore experimental data were only indicated by numbers and not by the names of the subjects. Therefore an experimenter effect concerning the physical hypotheses can be ruled out. However, this does not mean that the content of the psychological hypotheses was not influenced by the experimenter. It may well be that the experimenter subconsciously treated different subjects in a different way according to their personality characteristics. For instance it may well be that he was more friendly to ´extraverts´ than to ´introverts´ and so on. For this reason the experimental data concerning the psychological hypotheses might include to some extent effects of the experimenter-subject interaction.

The more specific hypotheses we have discussed so far did not contain specific assumptions on psychological variables. Since in the following hypotheses the psychological aspects are more important we will concentrate mainly on the variable DIM which is the only one which can clearly be interpreted in terms of the instruction given to the subjects.

The only hypothesis which does not refer to DIM is the hypothesis of pragmatic information on success and failure (20). In this hypothesis it is assumed that the variables PIS and PIG correlate with the sheep-goat variables SGFAK1-3. This, however, turned out not to be the case. Hence this hypothesis can be rejected. In general this seems to hold for the sheep-goat hypothesis (3) which apparently is not a good predictor variable for ´PK-success´. This seems to be in agreement with the findings of other researchers (see Dale 1946, Van de Castle 1958, Nash 1946, and Weis 1972). Although this is a negative result it may be of some theoretical interest because in ESP research the sheep-goat variable does seem to be a rather reliable predictor. Therefore one might argue that this difference is indicative for a difference between the ESP and PK process. This argument was also independently presented in a recent work of Roll (see Roll 1985).

The variable PRIOR1 (the confidence score), however, which could be perceived as a kind of sheep-goat state variable shows a positive correlation to the ´PK-success´ (DIM) for Schmidt as well as for Markow runs. Figure 4 reveals that even the detailed structure is similar. It seems that those subjects with extreme confidence scores are less successful than subjects with moderate scores. This was also my impression during the pilot study.

The difference between the ´sheep-goat´ state (confidence score) and trait (sheep-goat questionnaire) seems to be rather interesting and requires further research.

Another hypothesis yielding a result in contrast to our expectation is the externality hypothesis (4). It assumes that subjects who believe that their life depends on other powerful persons or on external fate should be more successful in a PK experiment than subjects who feel more self-confident. This hypothesis was formulated in this way because Bender et al. (1976) found that mainly ´external attributing´ persons reported about alleged spontaneous metal bendings. However, since externality is generally connected to the feeling of powerlessness this assumption seems rather implausible in the case of a PK experiment. Nevertheless the initial formulation (see Lay 1982) was maintained to avoid the criticism of selectivity. The data show only a rather weak relationship between the PK-success and the externality factors. IPC-factor 1 (externality) shows a negative correlation in the case of both Schmidt and Markow runs. This is formally in contrast to our hypothesis, but not entirely implausible.

In contrast to the ones discussed above, the remaining psychological hypotheses were formulated on a more reliable basis since the same questionnaires had been used in a previous multivariate PK experiment (see Weis 1972, Mischo & Weis 1973). Therefore the type and the tendency of the expected correlations between the FPI and EWL questionnaires and the variable DIM could be directly derived from this previous study. The hypotheses are indicated in the column H1/T of the correlation matrices (last two columns) and the results which gives the tendencies which were found in the experiment in the column TNDZ. Only for the Markow condition the predicted relationships are found (see also figure 4).

The extraversion hypothesis (5) assumes that subjects with a high score on the FPI-scales 5 (sociability), 6 (calmness), 10 (extraversion) and 12 (masculinity), will be more successful than their corresponding antagonists and those with high scores on scales 1 (nervousness), 3 (depressivity), 4 (excitability) and 8 (inhibition) will show negative correlations to ´PK-success´. Since these relationships were found (except for scale 10!) we can accept the extraversion hypothesis. A more detailed discussion of these correlations can be found in Lucadou 1986. Of special interest seems the masculinity scale, which shows the most prominent correlations (also for the Schmidt RNGs!). Independently from this study Roll (1986) argues that this psychological construct should be expected to be the most prominent predictor in PK experiments.

The hypothesis of high spirits (7) assumes that EWLFAK3 (high spirits) correlates positively with PK-success whereas EWLFAK1 (non active mood) and EWLFAK2 (depressed mood) correlate negatively with PK-success. In spite of the fact that only the factors EWLFAK1 and EWLFAK2 (see table 2) show the expected correlations for the Markow RNGs we accept the hypothesis because these correlations also show the predicted direction. However, the tendencies for the state variables (EWL variables) in general are much less clear since they show a rather non-linear dependency. This seems plausible in view of the underlying concept of state variables. They involve those psychological conditions (mood) which seem not stable in time. In fact many persons reported after the experiment that the experiment had changed their mood. This may be a reason for the observed non-linearity.

In the experiment of Weis the concept of frustration played an important role. He applied the concept in an experimental condition by introducing a (nearly unsolvable) puzzle task which the subjects had to solve before he or she was allowed to start the PK experiment. Since we did not like to introduce such conditions we used the concept of frustration simply as a variable which was measured before the experiment with the EWL-questionnaire. According to Weis´ results we hypothesized that frustrated subjects according to EWLFAK5 would be more successful that non frustrated ones. Again it turned out that this variable is a rather good predictor for the Markow RNG, but it is not only connected to the hit variables associated with F1 but mainly to the variance variables F5 (see table 2). In agreement with the results of Weis frustrated subjects appear to score better than non-frustrated subjects. However, they also ´produce´ more variance, which at the same time supports the variance hypothesis (see above). This result can be interpreted in different ways. One could argue that according to Stanford´s conformance behaviour model (Stanford 1974) frustrated subjects would have a certain ´need´ to reduce their frustration, for instance by performing a successful PK experiment. Or one could assume that frustrated subjects have a larger ´intentional distance´ toward their task similar to the ´release of effort effect´ (Rhine & Rhine 1943). This could also be applied as an argument for the increase in variance.

In any case the discussion of the relationships of the more specific psychological variables with the physical ones is rather speculative if one tries to ´understand´ it from psychological models, but at least the results do not seem to be implausible. It would be premature to try to interpret each of the ´discovered´ relationships in detail, since the considerable uncertainty of psychological constructs cannot be underestimated. On the other hand the surprising agreement of our results with the study of Weis and the more qualitative findings of other researchers should be taken into account if one tries to draw a final conclusion from this experiment even if this correspondence only holds for the Markow condition. In the moment it is impossible to find a reasonable ´explanation´ for this difference between the Schmidt- and the Markow condition. It might be a simple one, namely the lack of sensitivity, of the former, but even if we would know this we do not know why. In any case this is a rather astonishing result which requires further research. Intuitively, one would expect that the Schmidt RNG is more sensitive to PK because it has a larger variance.

7.7 Conclusions

If one considers all the different aspects of the evaluation of this experiment it seems justified to accept the PK hypothesis despite the fact that the acceptance cannot be based on a single p-value. This acceptance is mainly based on the meaningful correlations observed between psychological variables and physical variables and the difference these relationships exhibit under feedback and non-feedback conditions. The relevance of the feedback is strongly supported by the results of the special variable (DIM) which operationalizes the display-function to the subject. In seemingly contrast to the acceptance of the PK hypothesis stands the conclusion that no indication for a PK-signal (tracer) could be found.

Within the accuracy of the methods applied, no indication of a common PK-influence on the radio-active source or the five Geiger-Mueller tubes could be found nor was any special PK-signal (tracer) found which would allow to operationalize PK without making reference to psychological variables or conditions.

This seemingly paradoxical result casts serious doubts about the classical concept of PK as a real force or energy which is transferred from the subject to the PK-equipment. On the other hand it is in agreement with the assumptions of the OT´s which assume PK to be a non-local observer effect. Moreover, the data allow a decision between three competing OT models. Due to the different size and structure of the PK-effect on two types of RNGs (Schmidt and Markow) one central assumption of the Schmidt model seems to be refuted. The relevance of the variable DIM may furthermore indicate that Walker´s model has to be improved by the MPI. On the other hand the data seem not to support the IDS-model. The prominent role of the variable DIM further indicates that mainly the subjects are responsible for the observed effect, although it cannot be ruled out entirely that the experimenter and the whole experimental setting may also have influenced the psychological content of the PK-correlations.

These correlations indicate that extraverted, non-anxious, non-neurotic, non-depressive, non-inhibited and masculin subjects are more successful in a PK experiment than the corresponding antagonists. High spirits and frustration seem to be good state predictor variables.

One might criticize in the present study that the decision on the PK hypothesis cannot be expressed in one over-all p-value based on all different conditions and variables of the experiment. This would have required a much more homogeneous character of the variables studied. For instance, only those variables should have been included of which the distributions (under HO) are already known or are totally comparable for the two types of RNGs (if such is possible at all). Since, however, the primary aim of this experiment was to find a possible tracer-variable it was necessary to include ´unknown´ features. From this point of view, many of the findings of this experiment have mainly a heuristical character. Nevertheless it was decided not to restrict the number of variables for the final evaluation even after it had become clear that no tracer could be found, because this would have implied data selection on the basis of the data of the study.

Moreover, there is also a more fundamental argument which is often overlooked in parapsychology: An overall p-value can roughly be considered as a generalized measure for the strength of an effect or - in the case of parapsychology - of an anomaly. It contains ´functional information´ (see Lucadou 1987) about the effect. However, any effect or phenomenon also conveys ´structural information´ which is complementary to the former one. It cannot be expressed in a single value or function, it can only be expressed in ´pictures´ or ´matrices´ which exhibit the interrelationship between the elements as a pattern. This is the main reason for delivering the correlation matrices as the final result. Each element represents a p-value concerning a specific experimental question but until now we have no rule how to combine them into a final picture.

Normally, scientist do their experiments not ´in parallel´ but sequential. They study one question after the other selecting those which seem to be fruitful and neglecting others which does not seem promising. Finally, they develop an ´internal´ pattern of experimental results which represents the structural matrix of the interactions they investigate. This self-referential process is highly selective and does only work fruitfully if a certain theoretical prestructure is given. This prestructure, however, is still lacking in parapsychology. For this reason I propose parallel experimental processing (PEP) as a possible methodological approach to parapsychology.

ABSTRACT

Part III describes the physical hypotheses of the experiment, the evaluation methods, and the data reduction procedures. The reasons for the systematic bias observed in the RNGs is discussed. Some data from

the pilot study are described. As regards the evaluation of the main experiment among others four main hypotheses were tested:

1. PK (here defined as): There is a correlation between the physical process (variables) and psychological variables.
2. Feedback hypothesis: There is a difference in such correlations between feedback and non-feedback condition.
3. Generator hypothesis: There is no difference in such correlations between the two types of RNGs (Markow and Schmidt).
4. Tracer hypothesis: There is a specific physical process or variable indicating the influence of the PK effect per se which allows to distinguish the PK-process from stochastical behaviour, without having to include psychological variables. Especially the question was investigated whether PK has a possible direct influence on the radio-active source or a common influence on all five detectors.

The data support the PK and feedback hypotheses. The Markow´s RNG appears to be more sensitive to PK, which means that with the Markow RNG higher correlations with psychological variables are observed. Therefore the generator hypothesis 3 has to be rejected. Even though one can say that a PK effect could be demonstrated no indication for a tracer was found. With the given accuracy of measurement neither a common influence on the five detectors could be found nor variables showing a PK effect which could be distinguished from pure stochastical behaviour. Hence the tracer-hypothesis can also be refuted.

These findings raise a lot of theoretical problems concerning the so-called ´observational theories´. They are suggestive, however, for a system-theoretical model which starts from the assumption that ´allowed´ stochastical fluctuations of the psycho-physical system (process plus observer) can be triggered by the observer according to his or her (sub)conscious wishes. Feedback is essential for the triggering process. These correlations between psychological and physical variables are regarded as being non-local and they reflect the meaning (pragmatic information) of the display and the instruction given to the subjects. The assumption that the effect is due to precognition or intuitive data selection (IDS) is not supported.

Some fundamental considerations concerning such multivariate experiments are added which lead to the proposal of a parallel experimental processing (PEP) as a possible methodological approach in parapsychology.

Keywords: Psychokinesis (PK), observational theories (OT), psycho-physical systems, pragmatic information, intuitive data selection (IDS), model of pragmatic information (MPI), parallel experimental processing (PEP).

REFERENCES

Alcock, J. ´Parapsychology, science or magic? A psychological perspective´. Oxford, Pergamon Press, 1981.

Ballstaedt, C.Ch. ´Entwicklung und Evaluation eines Instruments zur Erfassung von Einstellungen zur Psychokinese´. Unpublished MA thesis, Freiburg, Psychologisches Institut, 1979.

Batcheldor, K.J. ´PK in sitter groups´. Psychoenergetic systems, 1979, 3, 77-93.

Bauer, E., Lucadou, W. v. ´Methoden und Ergebnisse der Psychokinese-Forschung´. In Condrau, G. (ed.) ´Transzendenz, Imagination und Kreativitaet. Die Psychologie des 20. Jahrhunderts´. Band XV, Zuerich: Kindler, 1979, 494-512.

Bender, H. ´Parapsychologie - ihre Ergebnisse und Probleme´. Hamburg, Fischer, 1976.

Bender, H. ´Neue Entwicklungen in der Spukforschung´. In Moser, F. ´Spuk - ein Raetsel der Menschheit´. Olten, Freiburg, Walter, 1977, 347-387.

Bender, H., Hampel, R. Kury, H., Wendland, S. ´Der ´Geller-Effekt´ - eine Interview- und Fragebogenuntersuchung. Teil II: Einstellungen, Psi-Informationsstand und ´okkultes Umfeld´ der Berichterstatter´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1976, 18, 217-241.

Betz, H.D. ´Unerklaerte Beeinflussung eines elektrischen Schaltkreises durch Silvio´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1978, 20, 47-50.

Bierman, D. Personal communication, 1982.

Blackmore S.J. ´Unreliability: Parapsychology´s only finding´. Paper presented at the 32nd International Conference of the Parapsychology Foundation, San Antonio, October 1983.

Broad, W., Wade, N. ´Betrug und Taeuschung in der Wissenschaft´. Basel, Boston, Stuttgart, Birkhaeuser, 1984.

Brunswick, E. ´Perception and the representative design of psychological experiments´. Berkeley, University of California Press, 1956.

Cattell, R.B. ´Personality: A systematical theoretical and factual study´. New York, McGraw-Hill, 1950.

Cattell, R.B. ´Personality and motivation: Structure and meaning´. New York, Harcourt Brace Jovanovich, 1957.

Claus, G., Ebner, H. ´Grundlagen der Statistik´. Thun, Frankfurt a.M., Harri Deutsch, 1979.

Dale, L. ´The psychokinetic effect. The first A.S.P.R. experiment´. J. A.S.P.R., 1946, 40, 123-151.

Ehrenstrasser, G. ´Stochastische Signale und ihre Anwendung´. Heidelberg, Huenthig UTB, 1974.

Eysenck, H.J. ´The structure of human personality´. London, 1953.

Eysenck, H.J. ´Personality and ESP´. J. S.P.R., 1967, 44, 55-71.

Fahrenberg, J. ´Probleme der Mehrebenen-Beschreibung und Prozessforschung´. Forschungsberichte des Psychologischen Instituts der Albert-Ludwigs-Universitaet Freiburg, Nr. 8, 1982.

Fahrenberg, J., Selg, H., Hampel, R. ´Das Freiburger Persoenlichkeitsinventar FPI´. Goettingen, Hogrefe (3.Aufl.), 1978.

Fishbein, M., Ajzen, I. ´Attitudes towards objects as predictors of single and multiple behavioral criteria´. Psychological Review, 1974, 81, 59-74.

Franz, M. L.v. ´Ein Beitrag zur Diskussion der Synchronizitaetshypothese C.G. Jungs´. In Bauer, E., Lucadou, W.v. (eds.) ´Spektrum der Parapsychologie´, Freiburg, Aurum, 1983, 94-104.

Hansel, C.E.M. ´ESP and parapsychology: A critical reevaluation´. Buffalo, New York, Prometheus Books, 1980.

Hasted, J.B. ´The metal-benders´. London, Boston and Henley, Routledge, Kegan Paul, 1981.

Hoevelmann, G.H. ´Are psi experiments repeatable? A conceptual framework for the discussion of repeatability´. E.J.P., 1984, 5, 3, 285-306.

Honorton, C. ´Replicability, Experimenter influence and parapsychology´. AAAS Paper, February 17, 1978.

Houtkooper, J. ´Observational theory: A research programme for paranormal phenomena´. Lisse, Swets and Zeitlinger, 1983.

Huber, K. ´Konstruktion und Bewertung eines Instruments zur Erfassung von Einstellungen zur aussersinnlichen Wahrnehmung´. Unpubished MA thesis, Freiburg, Psychologisches Institut, 1979.

Irwin, H.J. ´Psi and the mind´. Metuchen, N.J., Scarecrow Press, 1979.

Jahn, R.G. ´The persistent paradox of psychic phenomena: An engineering perspective´. Proceedings of the IEEE, 1981, 70, 2, 136-170.

Jahnke, W., Debus, G. ´Die Eigenschaftswoerterliste EWL. Eine mehrdimensionale Methode zur Beschreibung von Aspekten des Befindens´. Goettingen, Hogrefe, 1978.

Johnson, M., Haraldsson, E. ´The Defence Mechanism Test as a predictor of ESP scores: Icelandic studies IV and V´. J.o.P., 1984, 48, 185-200.

Kornwachs, K., Lucadou, W.v. ´Beitrag zur systemtheoretischen Untersuchung paranormaler Phaenomene´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1977, 19, 169-194.

Kornwachs, K., Lucadou, W.v. ´Psychokinesis and the Concept of Complexity´. Psychoenergetic systems, 1979, 3, 327-342.

Kornwachs, K. , Lucadou, W.v. ´Pragmatic information and nonclassical systems´. In R. Trappl (ed.) ´Cybernetics and systems research´, Amsterdam, New York, North Holland, 1982, 191-197.

Kornwachs, K., Lucadou, W.v. ´Pragmatic information as a nonclassical concept to describe cognitive processes´. Cognitive systems, 1985, 1, 1, 79-94.

Krampen, G. ´IPC-Fragebogen zu Kontrollueberzeugungen (´Locus of Control´)´. Goettingen, Hogrefe, 1981.

Kunzmann, H. ´Differenzierende Analyse der psychologischen Daten und Treffermessdaten in Zusammenhang mit einem multivariaten Psychokinese-Experiment´. Unpublished MA thesis, Freiburg, Psychologisches Institut, 1984.

Lay, B. ´Ein multivariates Psychokinese-Experiment´. Unpublished MA thesis, Freiburg, Psychologisches Institut, 1982.

Lay, B., Mischo, J. ´Das Sheep-Goat-Konstrukt in der parapsychologischen Forschung II´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie 1985, 27, 93-104.

Lichtenstein, S. Fischhoff, B., Phillips, L.D. ´Calibration of probabilities: The state of art to 1980´. In Kahneman, D., Slovic, P., Tversky, A. (eds.) ´Judgement under uncertainty: Heuristics and biases´, Cambridge, 1980.

Lucadou, W.v. ´Zum parapsychologischen Experiment - eine methodologische Skizze´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1974, 16, 57-62.

Lucadou, W.v. ´Vorschlag einer elektronischen Apparatur zur moeglichen ´Verstaerkung´ statistischer PK-Effekt´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1979, 21, 193-199.

Lucadou, W.v. ´What is wrong with the definiton of psi?´. E.J.P., 1984, 5, 261-283.

Lucadou, W.v. ´Experimentelle Untersuchungen zur Beeinflussbarkeit von stochastischen quantenphysikalischen Systemen durch den Beobachter´. Frankfurt, Haag + Herchen, 1986.

Lucadou, W.v. ´Keine Spur von Psi - Zusammenfassende Darstellung eines umfangreichen Psychokineseexperiments´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1986, 28, 169-197.

Lucadou, W.v. ´Non locality in complex systems´. Paper presented at the workshop ´Self-steering and cognition in complex systemes: Toward a new cybernetics´, May, 1987, Free University of Brussels (will be published).

Lucadou, W. v. ´The model of pragmatic information (MPI). Paper presented at the 30th Annual Convention of the Parapsychological Association at Edinburgh, August, 1987.

Lucadou, W.v., Kornwachs, K. ´Methodologische Betrachtung zur Beschreibung komplexer Systeme´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1974, 16, 148-157.

Lucadou, W.v., Kornwachs, K. ´Grundzuege einer Theorie paranormaler Phaenomene´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1975, 17, 73-87.

Lucadou, W.v., Kornwachs, K. ´Can quantum theory explain paranormal phenomena? In E. Morris ´Research in Parapsychology 1976, Metuchen, N.J., Scarecrow. 1977, 187-191.

Lucadou, W.v., Kornwachs, K. ´Development of the systemtheoretic approach to psychokinesis´. E.J.P., 1980, 3, 3, 297-314.

Lucadou, W.v., Kornwachs, K. ´On the limitations of psi´. Psychoenergetics, 1983, 5, 53-72. (Short version in W. G. Roll et.al. (eds.) ´Research in Parapsychology 1982´, Metuchen, N.J., London, Scarecrow, 1983, 85-89.

Lucadou, W.v., Kornwachs, K. ´Beyond reductionism - Reconsidered´. E.J.P., 1983, 4, 357-376.

May, E.C. ´Psychokinetic research at SRI´. In C.B. Scott Jones (ed.) ´Proceedings of a symposium on applications of anomalous phenomena´, Santa Barbara, 1984, 105-123.

May, E.C., Hubbard, G.S., Humphrey ´New evidence for interaction between quantum systems and human observers´. Paper presented at the Fourth Annual Meeting of the Society for Scientific Exploration at Charlottesville, 1985.

Mischel, W. ´Personality and assessment´. New York, Wiley, 1968.

Mischo, J. ´Neue Perspektiven parapsychologischer Forschung: Der Weg zum multivariaten Experiment´. In Bauer, E. (ed.) ´Psi und Psyche´, Stuttgart, Deutsche Verlags Anstalt, 1974.

Mischo, J. ´Das Sheep-Goat Konstrukt in der Parapsychologischen Forschung I´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1982, 23, 129-139.

Mischo, J., Lay, B. ´Das Sheep-Goat Konstrukt in der Parapsychologischen Forschung II´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1985, 27, 93-104.

Mischo, J., Weis, R. ´A pilot study on the relations between PK scores and personality variables´. In Roll, W.G., Morris, R.L., Morris, J.D. (eds.) ´Research in Parapsychology 1972´, Metuchen, N.J., Scarecrow Press, 1973, 21-23.

Mischo, J., Wittmann, W. ´Ein multivariates ASW-Experiment I - III´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1980 (a), 22, 23-50, 1981 (b), 23, 95-110, 27-57.

Moser, F. ´Spuk in neuer Sicht´. (1950) In Bender, H. (ed.) ´Parapsychologie´, Darmstadt, Wissenschaftliche Buchgesellschaft, 1966, 524-542.

Nash, C. B. ´Position effects in PK tests with twenty-four dice´. J.o.P., 1946, 10, 51-57.

Nash, C. B., Richards, A. ´Comparison of two distances in PK tests´. J.o.P. 1947, 11, 269-282.

Owen, I.M., Sparrow, M. ´Eine Gruppe erzeugt Philip´. Freiburg, Aurum, 1979.

Palmer, J. ´Attitudes and personality traits in experimental ESP research´. In Wolman, B.B. (ed.) ´Handbook of parapsychology´. New York, Van Nostrand Reinhold, 1977, 175-201.

Palmer, J. ´ESP research findings: The process approach´. In H.L. Edge, R.L. Morris, J. Palmer, J.H. Rush ´Foundations of Parapsychology´, London, Routledge & Kegan Paul, 1986, 184-222.

Radin, D.I., May, E.C. & Thomson, M.J. ´Psi experiments with random number generators: A meta-analysis´. Unpublished manuscript, 1986.

Rao, K.R. ´Psi und Persoenlichkeit´. In Beloff, J. (ed.) ´Neue Wege der Parapsychologie´, Freiburg, Walter, 1980, 104-123.

Rhine, L.E., Rhine, J.B. ´The psychokinetic effekt: I. The first experiment´, J.o.P. 1943, 7, 20-43.

Roll, W.G. ´Will personality and consciousness survive the death of the body?´. Unpublished manuscript, 1985.

Roll, W.G. ´Poltergeists´. In Wolman, B.B. (ed.) ´Handbook of parapsychology´, New York, Nostrand, 1977, 382-413.

Rorer, L.G., Widiger, T.A. ´Personality structure and assessment´. Annual review of psychology, 1983, 34, 431-463.

Rosenthal, R. "Interpersonal expectancies: The effects of the experimenter´s hypothesis´. In R. Rosenthal and R. Rosnow (eds.) ´Artifact in behavioral research´, New York, Academic Press, 1969, 181-277.

Rotter, J.B. ´Generalized expectancies for internal versus external control of reinforcement´. Psychological monographs, 1966, 80, (1. whole No. 609)

Rush, J.H. ´Problems and methods in psychokinesis research´. In Krippner, S. (ed.) ´Advances in parapsychological research. Vol. 1: Psychokinesis´, New York, London, Plenum Press, 1977, 15-78.

Sannwald, G. ´Beziehungen zwischen parapsychischen Spontanerlebnissen und Persoenlichkeitsmerkmalen´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1962, 6, 28-71.

Schmidt, H. ´A PK test with electronic equipment´. J.o.P., 1970, 34, 175-181.

Schmidt, H. ´PK test with a high speed random number generator´. J.o.P., 1973, 37, 105-118.

Schmidt, H. ´Comparison of PK action on two different random number generators´. J.o.P., 1974, 38, 47-55.

Schmidt, H. ´A logical consistent model of a world with psi interaction´. In Oteri, L. (ed.) ´Quantum physics and parapsychology´, New York, Parapsychology Foundation, 1975, 205-222.

Schmidt, H. ´A simple random number generator for use with minicomputers´, J. A.S.P.R., 1977, 71, 171-176.

Schmidt, H. ´Comparison of a teleological model with a quantum collapse model of psi´. J.o.P., 1984, 48, 261-276.

Schmidt, H. ´Addition effect for PK on prerecorded targets. J.o.P., 1985, 49, 229-244.

Spada, H. Personal communication, 1987.

Stanford, R.G. ´An experimentally testable model for spontaneous psi events´. J. A.S.P.R., 1974, 68, 34-57 and 321-356.

Thalbourne, M.A. ´Extraversion and the sheep-goat variable: A conceptual replication´. J. A.S.P.R., 1981, 75, 105-119.

Thalbourne, M.A., Haraldsson, E. ´Personality characteristics of sheep and goats´. Personality and individual differences, 1980, 1, 180-185.

Timm, U. ´Tornier und die Parapsychologie´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1979, 21, 99-120.

Timm, U. ´Methodologische Probleme bei der Planung und Auswertung differentieller Psi-Experimente´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1982, 23, 140-160.

Timm, U. ´Statistische Selektionsfehler in der Parapsychologie und in anderen empirischen Wissenschaften´. Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1983, 25, 195-229.

Van de Castle, R.L. ´An exploratory study of some personality correlates associated with PK-performance´. J. A.S.P.R., 1958, 52, 134-150.

Varvoglis, M.P. ´Goal-directed and observer-dependent PK: An evaluation of conformance theory and observational theory´. Research in parapsychology 1984, Metuchen, N.J., Scarecrow, 1985, 64-67.

Vassy, Z. ´Theoretical and methodological considerations on experiments with pseudorandom number sequences´. J.o.P. 1985, 49, 2, 127-153.

Walker, E.H. ´Foundation of paraphysical and parapsychological phenomena´. In Oteri, L. (ed.) ´Quantum physics and parapsychology´, New York, Parapsychology Foundation, 1975, 1-53.

Walker, E.H. ´Comparison of some theoretical predictions of Schmidt´s mathematical theory with Walker´s quantum mechanical theory of psi´. Journal of research in psi phenomena, 1977, 2, 54-70.

Walker, E.H. ´The quantum theory of psi phenomena´. Psychoenergetic systems, 1979, 3, 259-299.

Walker, E.H. ´A review of criticisms of the quantum mechanical theory of psi phenomena´. J.o.P., 1984, 48, 277-333.

Walker, E.H. ´A comparison of the intuitive data sorting and quantum mechanical observer theories´. Unpublished manuscript, 1987.

Weis, R. ´Frustration und Psychokinese (Ein experimenteller Ansatz zur Uberpruefung moeglicher Psychokinese und deren Beziehung zu Persoenlichkeitsmerkmalen)´. Unpublished MA thesis. Freiburg, Psychologisches Institut, 1972.

Weizsaecker, E.v. ´Erstmaligkeit und Bestaetigung als Komponenten der pragmatischen Information´. In Weizsaecker, E.v. (ed.) ´Offene Systeme I´, Stuttgart, Klett, 1974, 83-113.

Wittmann, W., Schmidt, J. ´Die Vorhersagbarkeit des Verhaltens aus Trait-Inventaren. Theoretische Grundlagen und empirische Ergebnisse mit dem Freiburger Persoenlichkeitsinventar (FPI)´. Forschungsberichte des Psychologischen Instituts der Albert-Ludwigs-Universitaet, Freiburg, 1983, (Nr. 10).

Editorial consultants: Volume 6, 1985-1987

The editors of the European Journal of Parapsychology wish to express their appreciation to the following referees for their assistance in evaluating articles submitted to the European Journal of Parapsychology.

E. Bauer (Freiburg, Germany); Dr. J. Beloff (Edinburgh, Scotland); Dr. D.J. Bierman (Amsterdam, The Netherlands); H.G. Boerenkamp (Utrecht, The Netherlands); Dr. W.G. Braud (San Antonio, U.S.A.); Dr. R.S. Broughton (Durham, U.S.A.); Dr. D.L. Delanoy (Edinburgh, Scotland); Dr. M. Dybvig (Trondheim, Norway); Dr. Hoyt Edge (Winter Park, U.S.A.); Dr. E. Haraldsson (Reykjavik, Iceland); C. Honorton (Princeton, U.S.A.); G. Hovelmann (Marburg, Germany); Dr. W. von Lucadou (Freiburg, Germany); Prof. R.L. Morris (Edinburgh, Scotland); Dr. H. Schmidt (San Antonio, U.S.A.)