EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 5

NUMBER 1

NOVEMBER 1983

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)

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.

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.

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.5, 1983

IN MEMORIAM OF PROFESSOR DR. J.M.J. KOOY (1902 - 1983)

G. Zorab

The physicist, mathematician and electrical engineer, Professor Dr. J.M.J. Kooy who in his later years became an internationally known expert on rocket techniques and space travel, came into contact with parapsychology in the 1930's. His interest in the subject was aroused by reading J.W. Dunne's book, 'An Experiment with Time' (1927). This book made a great impression on him, for he had himself experienced a number of precognitive dreams which had greatly puzzled him.

Following Dunne's example, Dr. Kooy took the trouble to write down all those dreams which he believed were of a precognitive nature. His dream studies forced him to wake up during the night more than once in order to note down a dream he had just experienced and which he considered of enough interest to be rescued from oblivion. His dreams were generally of a symbolic nature, and not congruent in all details with the future reality. Still, a few of Kooy's dreams were completely congruent with what would occur later on. Kooy continued these dream-experiments for three years before the last war. It was only in the course of the last few months of this period that he sent Dr. Tenhaeff copies of the contents of his dreams within 24 hours of having had them. By doing so, Kooy obtained a witness who could testify that he had read the dream contents several hours or days before the dream came true. Several of Kooy's precognitive dreams were highly remarkable, and these he published in the 'Tijdschrift voor Parapsychologie', together with his explanatory hypothesis based on the Einstein-Ouspensky four-dimensional time/space world. In that world, time is non-existent; what does exist could be termed the 'eternal present'.

In those days, Kooy lived in a village near the city of Utrecht where Tenhaeff and several other men of some renown who were interested in parapsychology and survival research also lived. Kooy came in contact with these men, who thought that precognition was such a wonderful phenomenon that it hardly could be possible. Although that was the general attitude, Kooy's experimental approach to the subject attracted those interested in parapsychology and it made them listen to his explanatory hypothesis of precognition and study it. There is little doubt that one of Kooy's great achievements was that he was one of the first to bring Einstein's Relativity Theory to bear on the paranormal phenomenon of precognition. He was an expert on the above mentioned theory and knew exactly what he was talking about.

It was especially after World War II that Kooy's name became known internationally. It was then that he was invited to lecture on the significance of the reality of precognition. Owing to his own experiences, Kooy himself had not the slightest doubt about precognition being a fact of nature, just as so many other facts we know about.

Kooy lectured not only in Holland, but also in France and the United States. In the latter country, Dr. J.B.Rhine, then the dean of parapsychology, invited him to Duke University (Durham) to lecture to his group of parapsychologists and discuss with them the possibility that precognition points to the existence of a four-dimensional space-time world. In such a Cosmos, in which time is nonexistent and therefore past, present and future are all one, it also has to be accepted that our so thoroughly entrenched dogma that the cause always precedes the effect is totally meaningless. The same applies to the concept of causality itself.

It cannot be said that Kooy's conceptions concerning the reality of our universe as a four-dimensional space/time continuum evoked much response among parapsychologists, scientists and philosophers. The latter may have regarded Kooy's ideas (which also were those of Einstein and Ouspensky) far too revolutionary to be accepted. On the other hand it is also possible that the rejection of and indifference to Kooy's conceptions were based on the fact that only a small number of men were inclined to accept the authenticity of precognitive phenomena.

Even in his own country, Kooy was often not taken seriously enough or sometimes misunderstood. I well remember that during the period after the last World War when I was much in Kooy's company and had become a fervent adherent of his view of life I published a book on precognition (1953). In this book I discussed Kooy's four-dimensional space/time continuum, pointing out that an explanation of that remarkable paranormal phenomenon might possibly be found by accepting a cosmos as described by Kooy. The above quoted point of view provoked a loud outcry from a number of prominent Dutch S.P.R Council members (Dr. Tenhaeff was one of them) who attacked me, claiming that I had rendered parapsychology a very bad service by writing such nonsense in my book on precognition!

Kooy was a very kind man who never tired of receiving visitors wanting to be informed about Einstein's Relativity Theory or such problems as space, time, consciousness, etc.. I, too, am very thankful for the lessons Kooy gave me. We used to sit together for hours on end, he talking and I just listening. Kooy was a man of many talents. He lectured and wrote in several modern languages. He even had a poetic vein !

Many people will miss him. I, for one, shall not forget my good talented friend Johan Kooy, sitting in his armchair, drawing circles and diagonal lines to show me how I had to imagine the four-dimensional world.

EUR.J.PARAPSYCHOL., V.5, 1983

STATE OF CONSCIOUSNESS DURING FEEDBACK: EXPLORING THE TWO STEP MODEL OF PSI

Pieter C. van der Sijde Pieter Wesseling
Research Institute for Psi Phenomena and Physics
Amsterdam

The purpose of this study is to introduce and explore a model of psi embedded in the Observational Theories (OT). The OT have the advantage of offering a unified psi-theory by reducing all categories of psi to PK (see e.g. Walker, 1975). According to a survey by Millar (1978), GESP is a special case of time-displacement PK in which the subject's earlier guessing behaviour plays the part of a random event generator.

We propose to label our description of psi the Two Step Model (TSM). There are two important episodes in an experiment during which a subject can play a crucial role: at the moment of making a guess and at the moment of feedback. Thus in the TSM we distinguish two steps:
- step one: generating a random event (i.e. the guess);
- step two: retro-active PK biasing the random event of step one (at the moment of feedback).

It has been assumed that there are certain favourable conditions for the occurrence of psi (Honorton, 1977; Shapin & Coly, 1978). In particular, research has been conducted to determine the most

----

This paper fulfils the publication policy of this journal

favourable state of consciousness of a subject for psi to occur. These process-oriented studies have only examined the importance of the state of consciousness of a subject during what we call within our TSM the first step, the 'guess'.

Within our TSM, the optimal condition for psi is specified by two sets of states of consciousness: one set corresponding to the optimal state for step one (the guess) and one set corresponding to the optimal state for step two (the feedback). The ideal research design would be a design which takes both steps into account and in which the states of consciousness in both stages are manipulated. It should be noted that the TSM only deals with the importance of the two episodes in the psi-process treated separately: nothing is said about the relation between the two episodes for an optimal psi result.

To our knowledge there are two previous studies which illustrate our TSM, although it is also being used by D.H. Weiner (personal communication).

Both studies examined the importance of the second step in the psi-process. The first attempt to separate the effects of the two steps was undertaken in a clairvoyance study that explored the effect of sleep deprivation on ESP (Bosga et al, 1980). Two subjects each did 32 runs, of which 16 were forced choice and 16 were free response. The forced choice data did not show any psi, but the free response data did show an effect. In the free response part, the protocol produced by the subject in the sleep deprived state (step one) was compared by the subject, while still in the sleep deprived state, with the target (step two). This procedure was followed in half of the free response trials; in the other half the subjects did not receive any feedback. The difference between the two conditions was significant at the .01 level (two-tailed).

In a study by Bierman (in press) the effect of meditation in a remote viewing setting was explored. There were two conditions in the experiment, meditation and non-meditation. The state of consciousness for the subject was the same for step one as for step two. In this study a displacement effect was found that could only be explained under the assumption that the meditation state was favourable for the first step (protocol production) but prevented retro-active PK when this state continued during feedback (step two).

Contrary to the classical approach, we also focused on the differential effect in the second step of the TSM, and this was done by manipulation of the state of consciousness in the second step. It was assumed that during the second step the subject performs retro-active PK on his more or less random guessing behaviour generated during the first episode of the experiment.

To manipulate the state of consciousness, a derivation of the Christos procedure (McIntosh, 1979) was used. We call our procedure the Christos Induction Method (CIM).

The CIM was used to manipulate the state of consciousness of a subject in step one in order to enhance the random-like behaviour of the brain. At the moment of feedback the subject was brought to a normal state of consciousness or left (as much as possible) in the state he was brought into by the CIM (the CIM-state).

In the present experiment two conditions were compared:

condition one:	step one - CIM-state	step two - CIM-state;
condition two:	step one - CIM-state	step two - normal state.

To bring the subjects from the CIM-state back to a normal state the cold pressor procedure was applied (Calloway & Dembo, 1958), in which the subject's feet are put in a footbath of cold water.

The hypothesis to be tested is the difference in effect of the two conditions, to be evaluated by the Mann-Whitney U test (MWU). As exploratory measures, the galvanic skin response (GSR) was recorded, and a rating scale (Palmer and Vassor, 1974) was filled in by the subjects before and after the experimental session.

METHOD AND PROCEDURE

Subjects

The number of subjects was 30. Two subjects who did not meet the criteria for the CIM (to be explained in the procedure) were replaced by two other subjects. All subjects were volunteers. After they agreed to participate, a letter was sent to them with information about the experiment.

Experimenters

In all cases the experimenters were the authors. The subject-handling was done by the first author and the equipment-handling by the second author.

Materials

The GSR was recorded on paper by a paper-recorder, and the mean GSR for each 15-second interval was recorded on a computer file. The two GSR electrodes were attached on two different fingers (middle and index) of the right hand. By means of a signaling system, it was possible to indicate on the computer file the different stages in the experiment, and also to put a mark on the computer file when the subject did not meet the criterion for the visualization part of the CIM (see Procedure). The mentation of the subjects was recorded on cassette and transcribed.

Targets and Randomization

The targets were taken from a target-pool consisting of 10 sets of 4 pictures. The target-pool is an existing pool constructed by members of the Parapsychology Laboratory of Utrecht. Beforehand, a list of targets was randomly generated by a parapsychologist visiting RIPP at the time of the experiment and who had no other ties with the experiment. The randomization was open-deck: the same target could be chosen for more than one subject. For each trial the target was put in an opaque envelope. A random sequence of conditions was also produced. The assignment to conditions was closed-deck (15 subjects in each condition). For each trial there was an envelope indicating the condition.

Procedure

Before the subject (S) arrived, the equipment was prepared, i.e., the computer, the GSR-apparatus, the footbath, and the cassette recorder. When S arrived there was first an informal talk about parapsychology and S's own paranormal experiences. Then S was briefed about what was expected of him in the experiment and about the equipment to be used (which was also described in the letter he received beforehand). The pre-session questionnaire was filled in next, after which the experiment started. The GSR electrodes were attached to the S's fingers. After that began the CIM, which consists of the following parts:

- S lies on a bed with bare feet and closed eyes;
- S's ankles and feet are rubbed for about one minute by the first experimenter (E1), while the second experimenter (E2) is opening the target envelope elsewhere;
- S's forehead is rubbed for about one minute by E1;
- visualization:
    - S is asked to visualize that he "grows" through his feet for about 5 centimeters and then back to "normal", and he does this several times;
    - S is asked to visualize that he grows through his head for about 5 centimeters and then back to normal, and he does that several times. (Several times means: till the experimenter gets the impression that S is 'growing' and 'shrinking' quite easily);
    - alternating growing exercises are requested for feet and head (after the 5 centimeter exercises, 10,20 and 40 centimeters): this is repeated until E1 concludes that it is quite easy for S to do;
    - S is asked to grow in both directions at the same time (through the head and through the feet) and also to start to 'fill up' the experimental-room ('growing in all directions');

At this point E1 decides whether or not S has met the criterion for the visualization part. The criterion is that when S shows any sign of visualization (it does not matter how small), the decision is made by E1 on the basis of his observation that S is a real subject in the experiment; if not, this is marked immediately on the computer file by E1 using the signaling system. The experimental procedure is continued, but the results are not taken into account. The procedure then continues as follows:

- S is asked to visualize and describe his own front door and is guided by E1 who asks him questions about it;
- S is asked to look at a reflecting surface in his mental picture of his front door and visualize on that surface an image of the target (for the first time);
- S is asked to look at, and walk around in his visualized neighbourhood

and tell E1 what he sees, guided by questions from E1;
- S is led in his mind to the nearest pool of water and is asked to visualize on the water's surface the target (for the second time); this is also guided by E1.

After the second visualization of the target, E1 opened the condition envelope which contained the message 'water' or 'no water', corresponding to 'normal state' or 'CIM state'. In the case of 'water', S was asked to open his eyes, sit up and put his feet in the water. At that moment E2 came in and showed S (and E1) the target. In the case of 'no water' (CIM-state), E1 signaled E2 to come in and hold the target in front of S and asked S to open his eyes; the first thing S saw was the target. The next thing was to discuss the subject's mentations and the target, how the recordings were made, and to ask S to fill in a post-session questionnaire.

E1 was always blind to the target, and neither experimenter had information on the other three pictures in the set.

Judging

The transcribed protocols produced by the Ss were sent to an independent judge after the experiment was finished. Each target belonged to a set of four pictures. An experienced judge matched the protocols with the pictures in the corresponding target-set. Rankings as well as ratings were performed. (Ranking: 1 - most resemblance, to 4 - least resemblance. Rating: 0 - no resemblance, to 100 - perfect resemblance)

RESULTS

The hypothesis - which concerned the difference between the two conditions - could not be confirmed (MWU: $z=1.62$, $p=.10$) . Evaluating the hypothesis using ratings gave a similar result (MWU: $z=1.62$, $p=.10$) (see table 1).

The potential range of the ratings was 0-100, with the actual range being 0-80. The mean rating was 24.53. The mean rank was 2.4 (MCE: 2.5), thus there was no overall psi in the experiment. The mean rank

TABLE 1
ESP scores in relation to conditions

	rankings *		ratings **	
	mean	st.deviation	mean	st.deviation
normal state	2.6	.99	17.07	23.37
C.I.M state	2.2	1.08	32.00	24.41
overall	2.4	1.04	24.53	24.68
expected	2.5	----	----	----
	MWU: z=1.62		MWU: z=1.62	

*) ranking: 1 - most resemblance
4 - least resemblance
**) rating : 0 - no resemblance
100 - perfect resemblance

in the normal feedback state was 2.6 (MCE: 2.5), and the mean rating was 17.07. The mean rank in the CIM feedback state was 2.2 (MCE: 2.5), and the mean rating was 32.00 (see table 1).

FURTHER EVALUATIONS

There was a post hoc decline effect in the ratings over sessions represented by a negative correlation between the session-number and the rating ($r=-.33$, $p<.05$); there was no decline effect in the rankings ($r=.17$, ns).

The questionnaire

Based on the data of the pre-session and post-session questionnaires, the Ss were divided into three groups:

Group I — Ss who felt more relaxed after the session than before;

Group II — Ss who felt more tense after the session than before;

Group III — Ss who felt the same before and after the session (indifferent).

Within each of these three groups, no differences were found between the two experimental conditions (see table 2).

TABLE 2
Evaluation of the hypothesis in the three groups constructed on the basis of the questionnaire

Group	number of Ss	the difference between the two experimental conditions
relaxed	15	MWU: z = .94
tensed	6	MWU: z = .87
indifferent	9	MWU: z = .41

Between the groups, there were no significant differences in ratings between Group I and Group II (MWU: $z=1.71$, $.05<p<.10$), between Group I and Group III (MWU: $z=1.70$, $.05<p<.10$), nor between Group II and Group III (MWU: $z=.53$, ns).

One of the questions on the post-session questionnaire was about the 'weirdness' of the experience. There was a suggestive correlation between weirdness and ranks ($r=-.24$, $p=.10$); no correlation was found between weirdness and ratings ($r=.17$, ns). No correlation was found between the length of the session and the ranks: $r=-.03$.

TABLE 3
Comparison of the three groups constructed on basis of the questionnaire

group	mean rating*	standard deviation
relaxed	$\langle x \rangle = 34.87$	s.d.: 26.84
tensed	$\langle x \rangle = 14.33$	s.d.: 19.97
indifferent	$\langle x \rangle = 14.11$	s.d.: 17.27
relaxed-tensed	MWU: z=1.71	$.05<p<.10$
relaxed-indifferent	MWU: z=1.70	$.05<p<.10$
tensed -indifferent	MWU: z= .53	ns.

*) rating: 0 - no resemblance, 100 - perfect resemblance

The GSR

From the reports of Ss who participated in the experiment, it appeared that some of them experienced relaxation. By means of the GSR we had an objective measurement of relaxation. The Ss were divided into two groups on the basis of the GSR: Ss who were able to relax during the initial part of the session and Ss who were not able to do so. The division into two groups was done by two independent judges, each of whom judged all 30 Ss on the basis of the GSR tracings. The judges agreed for 17 Ss, and for 13 Ss they differed in opinion. These 13 were left out of the analysis.

Within the two groups there were no significant differences between the two experimental conditions: the relaxed-GSR group MWU: z=.45, the tensed-GSR group MWU: z=.29.

DISCUSSION

The results of this experiment did not support our hypothesis. This could have been due to several factors; e.g., the CIM state might not have enhanced the random guessing behaviour of the Ss.

A better procedure to enhance random guessing behaviour might be relaxation. Relaxation has been found to be a very 'stable' psi-favourable state (Honorton, 1977), but a comparison of relaxed with tense or indifferent Ss was not significant in the present study. However, when we compared post hoc relaxed versus not relaxed (tense and indifferent pooled), as was done, e.g., by Gerber & Schmeidler (1957), we did find an effect, MWU: $z=2.88$, $p<.01$. This could mean that for experimental research it might be more interesting to investigate the dimension relaxed-not relaxed than the dimension relaxed-tense. This result holds for reported relaxation but could not be confirmed by objectively measured relaxation. There was no significant difference between objectively measured relaxation and tension (MWU: $z=.19$), nor between relaxed-not relaxed as measured objectively (MWU: $z=-.48$).

The questionnaire responses might have been contaminated by feedback of the target. Palmer & Lieberman (1975), for example, found that target feedback affected responses on the Betts Scale of Imagery.

The overall post hoc decline in the ratings could be explained by a judging effect: the judge who received no instructions about the judging-order judged the protocols starting with the protocol of the first session and ending with that of the last session. It could also have been a psychological type experimenter effect, since the experimenters received feedback/information on how the experiment was going in terms of results.

In a paper at the SPR-Conference in Bristol, Harley (1981) presented a study about ESP in dreams in which it appeared that there was a correlation between 'weirdness' and psi. In the present study this could not be confirmed.

Although our results did not support our hypothesis, we think that the TSM is a worthwhile contribution to the study of psi-phenomena, all the more because it fits within the OT. In the OT, the importance of feedback is stressed (e.g. Houtkooper et al, 1980): without feedback no psi. More than in the 'classical' OT, in the TSM the emphasis is put upon the state of consciousness during the two important episodes. The TSM opens a whole new line of research, because not only is the first step open for experimental manipulation, but also the second step. It is our opinion that the TSM will ultimately make a contribution to process-oriented psi research and we hope that the present study will promote discussion and, most of all, research to explore the TSM further.

ABSTRACT

Within the Observational Theories (OT) we introduce the Two Step Model (TSM). According to this model, there are two times at which a subject can use his psi: at the moment of making the guess (generating a random event) and at the moment of feedback (retro-active PK biasing the random event).

In the present study, the state of consciousness during feedback was manipulated. During the first step, the subjects were placed in a dreamlike state (CIM-state). At the moment of feedback, half of them were brought back to a normal state of consciousness to receive feedback, while the other half received the feedback in the CIM-state before they were brought back to a normal state of consciousness.

The hypothesis to be tested was the differential effect of the two conditions of feedback. The data, however, did not support the hypothesis. A post hoc decline effect in the ratings ($r=.33$, $p<.05$) was found, and also a post hoc effect of reported relaxation (MWU: $z=2.88$, $p<.01$).

ACKNOWLEDGMENT

We would like to thank Dick Bierman, Joop Houtkooper, Brian Millar and Debbi Weiner for their help and fruitful discussions. Our special thanks to our subjects and to the Research Institute for Psi Phenomena and Physics (RIPP): they made the experiment possible.

REFERENCES

Bierman, D.J. 'Remote viewing and meditation: The effect of state of consciousness during protocol production and during feedback', to be published.

Bosga, D, H. Smid and Bierman D.J. 'Feedback effects on clairvoyance in a sleep-deprived state', Paper presented at the PA, 1980, abstract in Roll, W.G. and Beloff, J. (eds), Research in Parapsychology 1980, Metuchen, N.J., Scarecrow Press, 1981.

Calloway and Dembo, 'Narrowed attention: Psychological phenomenon that accompanies certain physiological change', Arch. of Neuropsychiatry, 79, 1958, 74-90.

Gerber, R. and Schmeidler G.R. 'An investigation of relaxation and of acceptance of the experimental situation as related to ESP score in maternity patients', J.o.P., 21, 1957, 47-57.

Honorton, C. 'Parapsychology and altered states of consciousness', in Wolman, B.B. (ed), Handbook of Parapsychology, New York, Van Nostrand Reinhold, 1977.

Houtkooper, J.M., Andrews, R., Ganzevles, P.G.J., van der Sijde, P.C. 'A hierarchical model for the observational theories: A study of subject identity and feedback in repeated retroactive psychokinesis', E.J.P., 3, 1980, 221-246.

McIntosh, A.I. 'The "Christos" procedure: A novel ASC induction technique', Psychoenergetic Systems, 2, 1979, 377-392.

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

Palmer, J. and Lieberman, R. 'The influence of psychological set on ESP and Out-of-the-Body Experiences', J. A.S.P.R., 1975, 69, 193-210.

Sargent, C.L., Harley, T.A., Lane, J. and Radcliffe, K. 'Ganzfeld Psi-Optimization in relation to session duration', in Roll, W.G. and Beloff, J. (eds), Research in Parapsychology 1980, Metuchen, N.J., Scarecrow Press, 1981.

Shapin, B and Coly, L. (eds), Psi and states of Awareness, New York, Parapsychology Foundation, 1978.

Walker, E.H. 'Foundations of paraphysical and parapsychological phenomena', in Oteri, L. (ed), Quantum Physics and Parapsychology, New York, Parapsychology Foundation, 1975.

P.C. van der Sijde
Leusdenhof 304
1108 DP Amsterdam-Zuidoost
The Netherlands

P. Wesseling
T.H.Haismastraat 30
9251 AV Bergen
The Netherlands

EUR.J.PARAPSYCHOL., V.5, 1983

IS THERE A PARANORMAL (PRECOGNITIVE) INFLUENCE IN CERTAIN TYPES OF PERCEPTUAL SEQUENCES ? PART I

Holger Klintman
Lund University

STUDY 1
TIME-REVERSED INTERFERENCE WITH REACTION TIME

For a number of years the present author, as an experimental psychologist, has been studying interference phenomena in perceptual sequences. In a paradigm frequently used, a first stimulus (S1) is followed after a brief time interval - usually of the order of .5 to 3 sec - by a second stimulus (S2). Under these conditions, if the meaning of S1 in relation to S2 is varied, the reaction time (RT) to S2 typically changes due to the cognitive interference of S1 with the subject's identification of S2. Thus, for instance, if S1 is a word with the same meaning as S2, the RT to S2 may be shorter than if S1 is semantically non-associated.

Normally in such designs RT is recorded only for the response to the second stimulus in each sequence, but in some experiments I included reaction times to the first stimulus as well (to establish individual base-line RT levels). As expected, the results showed the presence of cognitive interferences, that is, the RT to S2 characteristically changed with the meaning of S1.

However, in the course of the data analysis an apparent anomaly revealed itself: in addition to the perception of S2 being affected by the presence of the preceding S1, there seemed to be a second type of 'interference', viz. one where S2 affected the perception of S1. This was indicated by the characteristic changes in the RT to S1 as a function of the meaning of the subsequent S2.

It proved extremely difficult to find a rational explication for the anomaly, and after repeatedly making similar informal observations I decided to investigate the nature of the relationship in a specific experiment.

I had at the time some general familiarity with experimental parapsychology and realized, that if in no way the changes in the RT to S1 could be shown to be dependent on some factor preceding or concomitant with the response to S1, an interpretation of the relationship in terms of ESP could not be ruled out. Specifically, one hypothesis would be that the subjects' perceptions and identifications of S1 was - by way of precognitive or 'time-reversed' interference (TRI) - affected by their identification of S2, the latter occurring about a second later. The class of ESP phenomena which comes closest to describing this type of relationship, then, would be precognition.

As is well known to the readers of this article, in experimental parapsychology the problem of precognition has for decades been a center of interest. From the pioneering experiments of Rhine (1938) in the thirties, the PK versus precognition controversy of the fifties and sixties (e.g. Morris, 1968), the linkage of ESP with differential psychology (e.g. Nash, 1966; Johnson & Kanthamani, 1967) to experiments on ESP in animals (e.g. Duval & Montredon, 1968) a host of reports have indicated the existence of the phenomenon.

Of the variety of methods employed, many have relyed on subjects' conscious selection of one of several alternative responses at time T(0) as indicative of, or otherwise associated with, some future event at time T(1) and in the absense of logical inference. The hypothesis was that this selection would to some extent be dependent upon or correlated with the subsequent occurrence of the future event.

An example of a method not using conscious selection of responses is a study by Stanford and Stio (1976) using changes in RT as an indicator of ESP in a free-association task. Similarly, in the present study the method used was strictly non-introspective, the basic dependent measure being simply the time needed for the subject to identify a stimulus word flashed on a screen.

METHOD

Procedure

Each stimulus sequence consisted of a color surface (S1) followed by a color name (S2), both relating to the colors red, green, blue or yellow and flashed in succession on a screen in front of the subject. The inter-stimulus interval was 850 ms and two contingencies were used: congruence (the two stimuli having similar meanings, e.g. blue-blue, green-green, etc.) and incongruence (the two stimuli having different meanings, e.g. blue-green, red-yellow, etc.) The subjects' task was to name as quickly as possible first S1, and then S2, the response to S1 always being fully completed before the onset of S2 (see figure 1).

In the total series of 36 trials (sequences) S1 followed a fixed series, whereas the contents of S2 were contingent on a fast alternating random number generator (RNG), with a frequency of about 10 c/s, based on instable relay circuits. The output of the RNG selected one of two slide projectors for the exposure of S2. Projector 1 always displayed congruent and Projector 2 incongruent S2 stimuli. In addition, prior to each trial, the experimenter threw a switch, cross-connecting the leads from the RNG output to the two S2 projectors every second trial. Thus a second source of variation was added to the RNG. The reason for this was to guard against any systematic bias of the RNG towards one or the other experimental conditions yielding an excess of congruent or of incongruent trials. In such a case, the switching would tend rather to alternate these conditions, while irregular RNG series would still retain their randomness. This procedure was used in order to secure the balance between the two conditions over the total series and yet fulfill the requirement of adequate randomness. Thus which condition would obtain in a sequence (congruence or incongruence) could not be inferred or known by either the subject or the experimenter before the actual exposure of S2.

Further, the experiment was carried out under double-blind conditions, i.e. neither the experimenter nor the subjects were

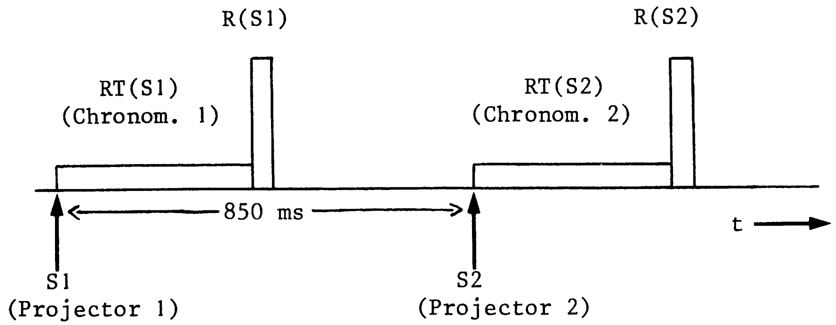

FIGURE 1
The stimulus-response sequences. S1 (color surface) is followed after 850 msec by S2 (color name). The subject's reaction times, RT(S), to these stimuli are measured, the response, R(S), being a verbal identification of each stimulus. The random generator selecting the S2 condition (congruence or incongruence) is continuously in operation.

informed about the nature of the hypothesis, or even that ESP was at all considered.

Subjects and experimenter

Twenty-eight first-year psychology students, aged 20 to 25 years, served as subjects. The experimenter was a younger colleague of the author (Mrs. Monica Henning, Lund).

Apparatus

The apparatus included a three-channel electronic timer, three slide projectors (one for S1 and two for congruent/incongruent S2), chronometers, speech amplifier (voice-key), and a random number generator.

Hypothesis

The hypothesis, as derived from earlier observations, was the following:

> If in a given subject there is a tendency to identify the second stimulus (S2) faster under congruent than under incongruent interstimulus conditions, then the same will be true for his identification of the first stimulus (S1).
>
> Conversely, if the identification of S2 is slower under congruent than under incongruent conditions, then the same will be true for the identification of S1.

RESULTS

Each subject's raw data consisted of the reaction times obtained from the 36 stimulus sequences (trials). From these data two differences were formed: D1 is the difference in RT to S1 between congruent and incongruent pairs, and D2 is the difference in RT to S2 between congruent and incongruent pairs. The following definitions were used for D1 and D2:

D1 = RT (c) - RT (i): The difference in RT to S1 between trial 1 and the first trial differing from trial 1 in congruence (congruent trial minus incongruent trial).

D2 = RT (c) - RT (i): The difference in RT to S2 between the first trial, not included in D1, and the first subsequent trial (not included in D1) with the same color and differing in congruence from the former (congruent trial minus incongruent trial).

Thus, the values of D1 and D2 were always derived from different pairs. This was necessary, since otherwise any RT baseline change acting upon both responses in a trial would tend to inflate trivially the correlations between D1 and D2. Also, the factors of color hue and color names were balanced. Since the time interval between S1 and S2 was always 850 msec, a few responses to S1 exceeding this value had to be excluded from the analysis.

Note also that in this design only four out of the 36 trials of each subject were included in the final analysis. This was a consequence of the random nature of the series in conjunction with the restrictions imposed upon trials selection by the requirements of balance in terms of color and of independence of observations (see definitions above).

According to the hypothesis, if in a subject the relation D2>0 were true (RT to S2 in congruent pairs greater than RT to S2 in incongruent pairs) then so would D1>0 (the same relation being true for RT to S1), while D2<0 would be associated with D1<0.

Table 1 summarizes the results. A significant relationship between the two variables is indicated (P=0.015, one-tailed) in the direction predicted by the hypothesis.

TABLE 1

Test of the relationship between D1 and D2. The entries are the numbers of the subjects with their respective D1/D2 combination.

	D1>0	D1<0
D2>0	14	5
D2<0	2	7

Fisher's exact P=0.015 (one tailed test)

Entered in the table are the number of subjects whose results were characterized by the respective D1/D2 relations. The group medians for D1 were 79 ms for the group with D1>0 and -47 ms for the group with D1<0.

The overall Spearman rank correlation between the D1 and D2 values for the 28 subjects was rho=0.5 ($p<.01$; df=26).

A test was also made for the effect of order between the conditions of congruence and incongruence as established by the RNG. No effect of this order on the relationship between D1 and D2 was found.

DISCUSSION

When evaluating the meaning of a result of this kind some caution is in order. An interpretation in terms of an ESP effect would be advisable only if any source of artifact - statistical or methodological - could be excluded.

In the present experiment, since in each trial the selection of the condition of congruence and incongruence was made electronically by the RNG, the selection taking place after the completion of the subject's response to S1 and physically independent thereof, the following potential sources of artifact were eliminated:

1 - The subject's 'reading' of conscious or unconscious cues from the experimenter.

2 - The biasing effect of a fixed series of stimulus combinations in previous trials on subject's stimulus expectancy in later trials.

As is well known (Rosenthal, 1969; Barber, 1976), in experimental work so-called subject-agent effects are always a potential source of error which may more or less skew the results systematically in the direction desired (or not desired!) by the person involved in the experiment. One necessary condition for such an influence, however, would seem to be that either the experimenter or the subject has some clue as to what would be required of him - in terms of performance during the experiment - in order to obtain a certain result. It is difficult to imagine that this could have been the case in the present experiment: first, since double-blind conditions were strictly upheld, second, since at the time of S1 no one could have inferred which condition would be chosen by the RNG, and third, since the skewing of RT to S1 in a direction required by the hypothesis would have required foreknowledge of not only the future state of the RNG, but also of the subject's performance in terms of RT under the contrasting condition (congruence or incongruence).

It should be mentioned that although the action of the RNG, when tested, roughly approximated what might be expected from chance in terms of the distribution of two alternative states, the device could probably not be considered an ideal source of random events. Specifically, it might be asked whether the incorporation of the alternator switch (see Procedure) might, in case of an RNG bias in the one or the other direction, be expected to yield above-chance alternation in the output. However, analysis reveals that even a bias of as much as 2 or 3 in favor of one output state would only slightly raise above .5 the probability of an alternation to occur (see Appendix 2). Therefore, and since certainly no bias of that order was present in the RNG used, the level of randomness could be considered satisfactory.

In view of these circumstances and of the fact that the experiment was prompted by earlier informal observations of similar relationships in daily experimental work, the hypothesis that precognitive interference played a role in producing the observed outcome was accepted, within the limits of confidence set by the statistical testing.

Apparently, the TRI effect indicated in the experiment need not have been of a magnitude leading to any conscious experience in the subjects; rather, the effect revealed itself in a slight temporal change in some phase of the perceptual/cognitive process of detecting and identifying the stimulus (S1). The change may be seen as a facilitating precognitive influence taking place, in some subjects, if the second event had the same meaning as the first ($D>0$), and in others if the two meanings were incongruent ($D<0$). Though this is not 'foreknowledge' with the usual connotation of being consciously experienced, still the crucial characteristics of precognition are present: the 'interference' of a later event with an earlier one within the same person, in the absence of logical inference.

Another point of interest is the role of individual differences. As we have seen, the corroboration of the hypothesis implied the presence of consistent differences with regard to the direction, or sign, of the time-change. In some of the studies reported below there is some evidence that this type of variation may be linked with differences on a general dimension of inhibition/facilitation.

In general, as a method, the detection of precognitive effects by means of changes in the temporal course of perceptual events seems a realistic supplement to the existing arsenal, its main advantage being high sensitivity and non-reliance on introspection and verbal report. As regards the two latter characteristics, apparently, a similar line of reasoning was followed by Hartwell (1978, 1979) in an investigation aimed at the detection of precognitive influence by means of the so-called Contingent Negative Variation (CNV) in the brain's electrical activity: since it is known that the CNV accompanies the subject's anticipation of imminent stimulation, it was hypothesized that under certain conditions precognitive anticipation (not necessarily conscious) might be indicated by the presence of CNV.

Thus, in terms of temporal patterning of the stimuli and, as we have seen, non-reliance on verbal report, the Hartwell study bears some resemblance to the present one. Although Hartwell did not report any significant psi effects, his method seems to the present author promising and well worth pursuing.

One limitation to the use of changes in the central statistics of reaction times, as in this study, has to do with the error variance associated with such measurements: it becomes increasingly difficult to detect effects smaller than 10-20 msec if the sample size is to be held within reasonable limits. This is a point of interest since one cannot rule out the possible occurrence of TRI shifts of a lower order in some individuals or constellations of events. As will be seen in the following studies, this difficulty may possibly be overcome by the use of temporal feedback loops.

STUDY II
CONCEPT VALIDATION WITHOUT THE USE OF AN RNG

At this point, having arrived at the conclusion that a paranormal influence might indeed be responsible for some of the observed aberrations, several different strategies for continued investigation could have been chosen. Perhaps the one which first comes to mind would be a direct replication, with no changes whatsoever in the design. Another would be an attempt at concept validation with a design using different stimulus material, and preferably, without the use of an RNG.

While direct replication might be the safer way to go, the advantages of successful concept validation (e.g. Hempel, 1966) were obvious: not only would the correlations observed in the special case of Study I receive confirmation, but the added information obtained in a different type of experiment might make possible more general conclusions about the nature of the phenomenon under scrutiny.

RATIONALE AND HYPOTHESIS

As a basis for Study II the central feature of the design of Study I was used: the succession of two cognitive events, separated by a brief interval. The general hypothesis was that under certain conditions the presence of the second event would paranormaly effect a change in the time course of the first event. The main differences between the two studies were the following (see also Method): 1) The stimuli used were digits rather than colors, and the associated two cognitive events were two successive comparisons of the two numbers on an odd/even basis. 2) No random generator was needed in the design, and 3) TRI was detected as changes in subjects' estimation (production) of a brief time interval, rather than in conventional RT to a stimulus.

The results of Study I indicated some major individual differences in the performance on the sequential task of identifying the two stimuli (S1 and S2). Thus, some subjects were facilitated (short RTs) in their identification of S2 by the preceding occurrence of a similar S1, whereas in other subjects a preceding non-similar stimulus, rather than a similar one, facilitated the second stimulus identification. Likewise, in the former group, there were indications that, in addition, the first task was facilitated, paranormaly, by the presence of a second similar task, while in the latter group of subjects such facilitation took place foremost in sequences of non-similar tasks. Influence of a first perceptual event on a second subsequent one has been reported and discussed by the author earlier, and individual differences in this respect have been found to correlate with suppressive/facilitative tendencies in the individual's negative aftereffect experience (Klintman, 1973). Specifically, in subjects with vivid aftereffect perception (measured in a separate test) the second of the two events separated by a brief time interval (<3 seconds) was found to be facilitated by the presence of a first event of equal meaning, while in subjects with less pronounced aftereffect experiences such facilitation took place if the two events had dissimilar meanings.

The fact that in Study I this 'forward' type of influence was accompanied by a 'backward' influence - presumably of a paranormal nature - made it reasonable to hypothesize a relationship between individual aftereffect perception characteristics and the specific stimulus conditions under which time-reversed interference would take place.

In the present experiment a negative afterimage duration test was used in which the cumulated time of the subject's afterimage perception was assessed (see Method).

Hypothesis

It was hypothesized that in a sequential task of making two successive numerical comparisons, a high and a low afterimage group would differ in the sense that in the high group the condition conducive to TRI of the second event with the first, would be two comparisons of equal meanings, while in the low group the corresponding condition would be non-equal meanings.

METHOD

Subjects and experimenter

Twenty-seven first-year university students, their ages ranging from 20 to 25 years, participated in the experiment. The experimenter, an experienced assistant to the author, was not the same person who served in study I (Mr. Bo Nystrom, Lund).

As before, double-blind conditions were used: neither the experimenter nor the subjects were informed in advance about the nature of the experiment. The information given was in very general terms, stating that the experiment was concerned with certain aspects of short-term memory and response time was to be measured.

The spreading effect

Before further discussing the method, brief mention should be made of the concept of TRI spreading. This is necessary for the understanding of the rationale of the design.

During earlier experimentation I had made some observations which indicated that TRI, once initiated, tended also to spread to other events closely related to those primarily involved.

To specify, let us consider an experimental set-up where the subject's task is to produce a time interval of the length of .5 to 2 sec on a digital counter by pressing two buttons in succession, one to start and one to stop the counting, and then to read on a display the number showing the exact length of the produced time interval. Now, if the perception of this number due to the subsequent occurrence of some other event was subjected to TRI, then the interference would also tend to spread to the associated cognitive processes involved in the subject's decisions to start and stop the timer. Assuming the resolution of the counter being high enough, the probability would be quite high that such secondary interference would change some of the numbers on the display.

Events and stimuli

The results of Study I suggested that sequences of perceptual/cognitive events such as the identification of a stimulus word or color may under certain conditions be accompanied by TRI, i.e., the second event upsetting the time course of the first. It seemed likely that such interference might not restrict itself to these relatively simple events but might indeed become even more pronounced if more complex and less stimulus bound events were used, such as for instance the comparing of two stimuli with respect to some abstract characteristic such as 'odd/even'.

With this in view, the events chosen for Study II were two successive comparisons of two printed one-digit numbers with respect to their being odd or even. Thus the events could take on one of two meanings: either 'equal' (the two numbers both being odd or even) or nonequal (the numbers differing with respect to the odd/even property). Furthermore, the results of Study I suggested the possibility that in one group of subjects ('facilitators') a necessary condition for TRI to occur in a sequence of two events was that the events have similar meanings, while in other subjects ('inhibitors') the condition would rather be dissimilar or incongruent meanings. In our present case, the former condition would be fulfilled in 'equal/equal' and 'nonequal/nonequal' sequences, the latter in 'equal/nonequal' and 'nonequal/equal' sequences.

Procedure

Each subject was given a separate afterimage test 10 minutes after the conclusion of the main experiment. The fixation stimulus was a light rectangle 15 cm wide and 6 cm tall, exposed on a semi-transparent screen, the light intensity being 8 lux (reflected light), as measured at 10 cm in front of the screen, and the background light in the room 4.5 lux, as measured in the same place. The measurements were made with a LUNA SIX lux-meter without a filter.

The subject was seated at a distance of 100 cm in front of the screen and was told to fixate on the rectangle (see Instruction, Appendix 1). The fixation time was 40 seconds, after which the projector was turned off. The subject's task was then, while still fixating on the same spot on the screen, to report to the experimenter every time a negative afterimage appeared or disappeared over a period of 100 seconds. The cumulated time for each such report was recorded by the experimenter. In the main experiment, the subjects' task (see Instruction, Appendix 1) was first to produce an estimated 1 to 2 sec. interval on an electronic counter and check the resulting rightmost digit on the counter display (time unit 1 microsec.); then to check the one-digit number printed on a card and compare the two numbers with respect to their being odd or even (is first comparison or event: response 'equal' or 'nonequal'); then, after three seconds (signalled by the experimenter), to produce another digit on the counter and compare it with the number printed on a second card, likewise with regard to the odd-even property (is second comparison or event: response 'equal' or 'nonequal'). The numbers on the cards were randomized and unknown to experimenter and subject until actually checked by the latter. Two series, each including five trials, were given (series a and series b) with a short break of 1 minute between them. The time between the trials was about 15 sec. The subject's responses were scored and a Typical Response Pattern (TRP) obtained (is the most frequent response combination: 'equal/equal', 'equal/nonequal', 'nonequal/equal' or 'nonequal/nonequal'), scored in the respective series.

The randomization of the cards was carried out prior to the experiment by a colleague of the experimenter. A randomnumber table was used in selecting the order of the cards. Since a reasonable balance between odd and even numbers was desirable, as well as the representation of all numbers between 0 and 9, several random number series had to be examined until these requirements were fulfilled.

During each session the experimenter kept a record of the numbers displayed by the timer and read by the subject. After the experiment but prior to the data analyses, this record was checked against the numbers in the card series and compared with the subjects' records of the results of the comparisons. If in a trial a discrepancy was found, then that trial was stricken from the data. This led to a data loss of about 5% due to subjects' recording errors. Figure 2 gives an overview of the temporal pattern of the experiment.

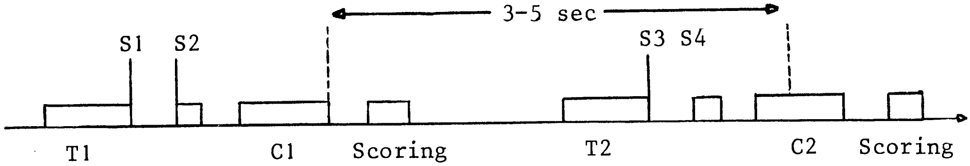

FIGURE 2

Temporal patterning of the experiment. T1 and T2 are the time intervals produced by the subjects, S1 and S3 the respective resulting one-digit numbers. S2 and S4 are the numbers on the cards of sets 1 and 2 checked by the subject at these times, and C1 and C2 are the comparisons of S1 to S2 and S3 to S4, respectively. Scoring of the outcome of these comparisons is also indicated.

Selective temporal feedback condition

The concept of temporal feedback will be discussed in some detail in Part II. In the present context it suffice to mention briefly its role for the formulation of the hypotheses.

Hypothesizing a spreading effect made it possible to build into the design a selective temporal feedback condition by letting the first number in each comparison be a measure (function) of the time interval produced by the subject: if we assume (see hypothesis) that in a given subject the condition conducive to TRI of the second event (comparison) with the first is that the events be identical, then, under this condition, the spreading effect would upset the timing of the associated decisions to start and stop the timer, frequently (P=.5) causing a change in the odd-even characteristic of the measurement (produced time length). All in all, provided the time shifts were greater than the time base setting of the counter (in this case 1 microsec.), the probability of actually observing identical responses would be reduced from a theoretical .5 to .25.

Statistically this may be likened to the case of tossing two unbiased coins: the probability (P) of obtaining non-identity (one head and one tail) in a first toss would equal the probability of obtaining identity (two heads or two tails), both being P=.5. Now lets stipulate that any non-identical outcome of a first throw be accepted as the final result, whereas in the case of identity the first coin is thrown a second time, again with a probability P=.5 of obtaining identity and P=.5 of obtaining non-identity. Apparently then, the P for obtaining identity in the final result would become P=.5*.5=.25, and that of non-identity P=.5+.5*.5=.75.

Conversely, if in a subject the condition necessary for TRI were two contrasting or non-identical events, the probability of observing non-identical responses would reduce from .5 to .25.

According to the hypothesis, then, subjects with a relatively strong tendency to facilitate internal perceptual/cognitive events (high afterimage scores) will show a greater proportion of non-identical response pairs than subjects with a tendency to inhibit such events (low afterimage scores).

RESULTS

Under the null-hypothesis, the frequency proportions of either of the four possible Typical Response Patterns (TPR) would be .25, the same for each of the two afterimage groups. Table 2 gives the distribution of TRPs for the combined two series of 5 + 5 trials. The numbers entered are subjects, each belonging to one of the two afterimage groups and having a TRP characterized by the two comparisons having either congruent or incongruent meanings. (Thus, for example, if the two comparisons yielded 'equality' and 'non-equality', respectively, then the trial would be entered under 'Incongruent', whereas 'non-equality' and 'non-equality' would be entered as 'Congruent'). As in subsequent analyses, subjects falling on the median in the AI test were excluded, as were those with a tied TRP (the number of congruent trials equals the number of incongruent trials).

TABLE 2
The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) of congruence (C) and incongruence (I) for the combined series of a and b (all trials).
Study II.

	TRP		
	C	I	=
AI>Median	4	9	(1)
AI<Median	6	5	-

Fisher's exact P>.05
Median(AI) 51 seconds

The weak contrasts found in table 2 fall above the 10% level of significance.

In a further analysis the data were partitioned with respect to 1) series (5 + 5 trials) and 2) the contents of the first comparison (C1). Tables 3 and 4 give the results. As in table 2 the entries are subjects belonging to either the high or the low afterimage group and characterized by one of the four possible TRPs. Table 3 indicates a statistically significant relationship in the predicted direction (p=0.029, one-tailed) between afterimage group and type of TRP. Thus, in trials where the first comparison yielded non-equality (odd/even or even/odd), the TRPs of subjects in the high group tended to be incongruent rather than congruent, while the TRPs of low group subjects were more often congruent. This relationship was more pronounced in series a than in series b. In contrast, no similar relationship is found in trials where the first comparison yielded equality (table 4).

DISCUSSION

Although the results indicated that under certain conditions a relationship exists between facilitative/inhibitive behavior in the afterimage perception test on the one hand and Typical Response Patterns on the other hand, further interpretation was suspended awaiting cross-validation (Studies III and IV).

STUDY III
CROSS-VALIDATION OF STUDY II

In Study III the experiment of Study II was repeated under identical conditions. The subjects were 22 students from the same category as in the previous studies. While the basic hypothesis was the same, the results of Study II allowed some further specifications. Thus it was predicted that the relationship observed in Study II between afterimage group and Typical Response Pattern would obtain only when the first comparison yielded nonequality, and then foremost in series a (the first half of the experiment, trials 1 - 5).

In table 5 are given the TRP distributions for all trials combined. It may be noted in this material that again the high group tended towards incongruent TRPs and the low group towards congruent ones. Tables 6 and 7 give the results separately for the two series (5 + 5

TABLE 3

The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded nonequality (ne) and C2 nonequality or equality (e). Study II.

	TRP ne /e 1 2	TRP ne /ne 1 2	=	TRP ne /e 1 2	TRP ne /ne 1 2	=
AI>Median	9	3	(1)	8	3	(2)
AI<Median	3	8	(1)	3	5	(4)

(a) Fisher's exact P=.029
Median(AI)=51 sec. Series a.

(b) Fisher's exact P=.144
Median(AI)=51 sec. Series b.

TABLE 4

The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded equality (e) and C2 equality or nonequality (ne). Study II.

	TRP e /ne 1 2	TRP e /e 1 2	=	TRP e /ne 1 2	TRP e /e 1 2	=
AI>Median	4	3	(6)	7	3	(3)
AI<Median	5	4	(3)	4	4	(4)

(a) Fisher's exact P>.20
Median(AI)=51 sec. Series a.

(b) Fisher's exact P>.20
Median(AI)=51 sec. Series b.

TABLE 5
The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) of congruence (C) and incongruence (I) for the combined series of a and b (all trials). Study III.

	TRP		
	C	I	=
AI>Median	2	6	(1)
AI<Median	6	2	(2)

Fisher's exact P=.066
Median(AI)=59 sec.

trials) and for trials were the first comparison yielded non-equality (odd/even, even/odd) and equality (odd/odd, even/even), respectively. As indicated in the tables 6 and 7, the results were very similar to those obtained in Study II. Thus, again a significant relationship was found between afterimage duration and TRP in series a when the first comparison yielded non-equality (p=0.004, one-tailed).

RESULTS OF STUDIES II AND III

Since Study III was a direct replication of Study II, a combination of the results in joint tables will give a clearer picture of the relationships involved.

Table 8 summarizes the overall undifferentiated results. The weak contrasts of the separate studies combined reach a level of statistical significance (p<.05, one-tailed). Note that the direction of the relationship is in accordance with the hypothesis.

TABLE 6

The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded nonequality (ne) and C2 nonequality or equality (e). Study III.

	TRP			TRP		
	ne /e	ne /ne	=	ne /e	ne /ne	=
	1 2	1 2		1 2	1 2	
AI>Median	6	1	(2)	5	2	(2)
AI<Median	0	6	(4)	4	4	(2)

(a) Fisher's exact P=.0041
Median(AI)=59 sec. Series a.

(b) Fisher's exact P>.20
Median(AI)=59 sec. Series b.

TABLE 7

The relation between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded equality (e) and C2 equality or nonequality (ne). Study III.

	TRP			TRP		
	e /ne	e /e	=	e /ne	e /e	=
	1 a	1 a		1 2	1 2	
AI>Median	2	3	(4)	2	4	(3)
AI<Median	2	6	(2)	4	4	(2)

(a) Fisher's exact P>.20
Median(AI)=59 sec. Series a.

(b) Fisher's exact P>.20
MD(AI)=59 sec. Series b.

TABLE 8
The relationship between negative afterimage duration (AI) and Typical Response Patterns (TRP) of congruence (C) and incongruence (I) for the combined series of a and b (all trials). Studies II+III.

	TRP		
	C	I	=
AI>Median	6	15	(2)
AI<Median	12	7	(2)

Fisher's exact P=.030
Median(AI) 55 sec.

In tables 9 and 10 the differentiated results are summarized.

STUDY IV
A FOLLOW-UP STUDY TO DETERMINE RE-TEST STABILITY

The purpose of this experiment was to study after 6 months the re-test stability of the results obtained in Studies II and III. However, for practical reasons only a limited number of subjects from the original experiment were available for a retest, most of them belonging to the high group in the afterimage test. It was therefore decided that the experiment would be restricted to this group.

The conditions of the experiment were the same as in the earlier two studies with some minor exceptions. One uninterrupted series of 7 trials was used. Also, this time the author served as experimenter.

In accordance with the earlier studies it was predicted that in the (high) group the observed Typical Response Pattern would be

TABLE 9

The relation between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded nonequality (ne) and C2 nonequality or equality (e). Studies II+III.

	TRP ne_1/e_2	TRP ne_1/ne_2	=
AI>Median	13	4	(3)
AI<Median	3	14	(5)

(a) Fisher's exact P=.00078
Median(AI)=55 sec. Series a.

	TRP ne_1/e_2	TRP ne_1/ne_2	=
AI>Median	13	5	(4)
AI<Median	7	9	(6)

(b) Fisher's exact P=.091
Median(AI)=55 sec. Series b.

TABLE 10

The relation between negative afterimage duration (AI) and Typical Response Patterns (TRP) for series a and b when C1 yielded equality (e) and C2 equality or nonequality (ne). Studies II+III.

	TRP e_1/ne_a	TRP e_1/e_a	=
AI>Median	6	6	(7)
AI<Median	7	10	(5)

(a) Fisher's exact P=.20
Median(AI)=55 sec. Series a.

	TRP e_1/ne_2	TRP e_1/e_2	=
AI>Median	9	7	(6)
AI<Median	8	8	(6)

(b) Fisher's exact P>.20
MD(AI)=55 sec. Series b.

predominantly non-identical and that this would be so only in the case of the first comparison yielding non-equality. The results are shown in table 11.

TABLE 11
Frequency distributions of Typical Response Patterns (TRP) when C1 yielded nonequality and C2 nonequality or equality and when C1 yielded equality and C2 equality or nonequality. Study IV.

TRP			TRP		
ne_1/e_2	ne_1/ne_2	=	e_1/ne_2	e_1/e_2	=
9	1	(2)	5	4	(3)
(a) P=.01	(Binomial test)		(b) P=.20	(Binomial test)	

As indicated in table 11 the results of the previous studies were replicated. The group displayed a greater proportion (p=.011, one-tailed) of non-identical TRPs when the first of the two successive comparisons contained a non-equality. There was thus evidence of some amount of re-test stability in the frequency contrasts resulting from this type of experiment.

CONCLUSIONS OF STUDIES II - IV

The results of studies III and IV (replications) were fully consistent with those of Study II. This indicated the presence of a systematic source of variation which under some specific conditions characteristically affects the subjects' timing of the decisions involved in the production of the time interval.

Before venturing any further interpretation of the results, let us examine some potential sources of artifact. First, as is well known (Rosenthal, 1969; Barber, 1976), in experimental work so-called subject-agent effects are always a potential source of error which may more or less skew the results systematically in the direction desired (or not desired) by the persons involved in the experiment.

Could, then, the results have been manipulated or skewed by the subject and/or the experimenter? Since the TRP was formed by the combination of two successive comparisons, C1 and C2, and since neither subject nor experimenter had any way of inferring the outcome of C2 at the time of C1, the manipulation of the resulting TRP would of necessity have had to involve the manipulation of one or both of the numbers constituting C2. In the design used in Studies II - IV, the two numbers compared in C2 were one produced by the subject on the digital counter and one printed on the next card in set 2. Since the series of numbers on the set 2 cards were fixed, the only way the subject (or the experimenter) could have affected the final outcome would be by changing the length of the time interval produced by the subject.

However, it seems extremely unlikely that such control of the timing of the interval would be possible: the units, repeating a series from 0 to 9 over and over again during the interval, each had a length of 1 millionth of a second. To control the outcome, the subject would have had to stop the timer at a digit (odd or even) yielding the 'desired' C1/C2 combination. Further, he would have had to do this without any information from the counter's display window, since the latter was covered by a screen during the whole of the interval. No evidence exists in the relevant literature of such extraordinary precision in human responses. But even if we assume such an improbable ability on the part of the subject, to produce a TRP attuned to the hypothesis he would still have had to take into account his relative standing on the afterimage test given after the completion of the TRI part of the experiment.

For the experimenter to manipulate the TRP outcome, the same is true as was for the subjects: he would, by means of some interference with the subject's performance, have had to bring about a controlled change in the length of the time interval at the basis for comparison C2, a task which, as we have seen, hardly lies within the reach of human abilities.

To continue our search for artifacts, could the results have been biased by accidental or voluntary faulty reports by the subject of the number produced on the counter? If in such reports the deviations from the actual number produced were unsystematic (random), this would have decreased rather than increased the probability of the observed relationships. As for the possibility of systematic deviations (for example, excess of 'odd' or 'even' responses), none were found when tested for.

In view of the fact that, in addition, double-blind conditions were applied throughout the series, it seems reasonable to conclude that the likelihood that any bias on the part of either experimenter or subjects might have produced the results appears to be negligable.

Next, it must be asked whether defects in the apparatus, lack of precision, or systematic time errors could have biased the outcome in the observed direction. Tests were run on the odd/even frequency distribution of the microsecond digit of the counter and also on its possible tendency to repeat a digit just displayed, but no such effects were found. And indeed it is hard to imagine how any technical deficiency could have correlated the resulting odd/even distributions of the displayed digits with the subjects' afterimage duration scores.

Consequently, it is difficult to explain the results in terms of ordinary psychological or technical artifacts. The rejection of these potential artifacts as responsible for the observed effect suggests that an interpretation in terms of precognitive (time-reversed) interference might indeed have some justification: support was found for hypotheses concerning:

(1) TRI. Thus in 'facilitators' (afterimage high group) TRI was found in sequences of equal events, and in 'inhibitors' (afterimage low group) in sequences of non-equal events.

However, consistently over the three experiments, the effect was present only in sequences beginning with a non-equality, foremost in the first half of the experiment (series a, trials 1 to 5). One difference between an equal match and a non-equal match may be that the latter gives rise to the stronger and more durable cognitive residual activity in the afterphase of the initial perception. This may have been of consequence in the present design, where the second match took place about 3 - 4 seconds later. It seems plausible that the state of this residual activity at the moment of the second match is an important factor for the occurrence of TRI. If so, the fact that the effect tended to decline after the first five trials might be understood as a gradual reduction, due to habituation, in the level of cognitive afterphase activity set up by the first match. A possible test of this hypothesis would be to repeat the experiment using a substantially shorter time interval between the two matches. Under these conditions, at the time of the second match, the level of residual activity following the first might be high enough to trigger TRI even in sequences beginning with an equal match. Also, the decline effect might be somewhat less pronounced over the series. These special conditions will be further discussed in Part II.

(2) the existence of a spreading effect.

(3) the possible occurrence of selective temporal feedback.

If we compare the results of Studies II, III and IV with those of Study I, a high degree of inter-experiment stability may be noted. This, in conjunction with the fact that in these experiments two different techniques were used as well as different types of stimulus events, suggests that TRI tends to be present generally in such perceptual/cognitive sequences where the two events are separated in time by 3 seconds or less. Also, since several different experimenters were used, there is some evidence that the results are not dependent upon any very specific personality characteristics on the part of the experimenter. Thus it is possible that these designs may have a fairly high degree of replicability.

In Part II a fifth study will be reported, dealing with the extension of the time range of TRI beyond the 3 sec interval. A heuristic model for the description of the causal relations involved will also be discussed. Finally, general conclusions from the six studies will be listed and some pertinent problems outlined as a basis for further investigation.

ABSTRACT

Four experimental studies of precognitive or Time-Reversed Interference (TRI) in perceptual sequences of two successive stimuli (S1 and S2) are reported. The samples consisted of unselected university students, the sample sizes varying between 12 and 28 subjects. In Study I the dependent measure was the reaction time (RT)

to color names and color surfaces, with an interstimulus interval of 850 msec. The results indicated that in addition to the (forward) interference of S1 with the RT to S2, there was also interference of S2 with the RT to S1. In studies II - IV the findings of Study I were used as a basis for a series of (concept) validation experiments using a different detection technique and stimulus-response paradigm. As a whole, the results clearly pointed to the presence of TRI and gave some support to an interpretation in terms of temporal feedback loops. Also, the individual's perceptual/cognitive style (facilitative vs inhibitive) played a major role for the prediction of the conditions under which precognitive interference would take place.

REFERENCES

Barber, T.X. 'Pitfalls in human research: Ten pivotal points', Elmsford, New York: Pergamon Press, 1976.

Hartwell, J.W. 'Contingent negative variation as an index of precognitive information', E.J.P., 2, 1978, 83-103.

Hartwell, J.W. 'An extension to the CNV study and an evaluation', E.J.P., 2, 1979, 358-364.

Duval, P., Montredon, E. 'ESP experiments with mice', J.o.P., 32, 1968, 39-55.

Hempel, C. 'Vetenskapteori', New Jersey: Prentice Hall, 1966, 40-45.

Humprey, B.M. 'Introversion-extraversion ratings in relation to scores in ESP tests', J.o.P., 15, 1951, 252-262.

Johnson, M., Kanthamani, B.K. 'The defence mechanism test as a predictor of ESP scoring direction', J.o.P., 31 1967, 99-110.

Klintman, H. 'Cognitive interference in the afterphase of perception as related to spiral aftereffect duration', Psycholog. Res. Bull., Lund University, 18, No 9, 1973.

Morris, R.L. 'Obtaining non-random entry points: A complex psi task', In J.B. Rhine and Brier (eds.), Parapsychology Today, New York, Citadel Press, 1968.

Nash, C.B. 'Relation between ESP scoring and the Minnesota multiphasic personality inventory', J. A.S.P.R., 60, 1966, 56-62.

Rhine, J.B. 'Experiments bearing on the precognition hypothesis', J.o.P., 2, 1938, 38-54.

Rosenthal, R. 'Artifact in behavioral science', New York: Academic Press, 1969.

Stanford, R.G., Stio, A. 'A study of associative mediation in psi-mediated instrumental response', J. A.S.P.R., 70, 1976, 55-64.

Holger Klintman
Department of Psychology
Lund University
Paradisgat 5
S-22350 Lund
Sweden

APPENDIX I

Instructions

STUDY I

"The purpose of this experiment is to measure your reaction time to colors and color names. That is, we will try to determine how fast you can name a color and then read a color name. The colors involved are red, green, yellow, and blue. In each display you will be presented first with a color surface and then with a color name. The time interval between the surface and the name will be a little less than a second. The meanings of the two stimuli will not always agree. For example, the color may be red and the word 'blue', or the color green and the word 'yellow', etc., while in other instances the two may agree, the word being the name of the color. Your task will be to first name the color surface as quickly as possible, and then the word following the color surface. It is important that you react as fast as possible. Try not to think about any specific color before each display. Speak loudly and clearly into the microphone..."

STUDIES II - IV

"The device in front of you on the table is a digital counter. If you press this button [indicating] the counter starts, and when you press this one the counter stops. You will then see on the display a number between 0 and 9. The procedure will be as follows. When I say 'digit', you will start the counter and then stop it after about 1 or 2 seconds. During this time interval the display window will be covered by the cardboard screen, like this [demonstrating], but as soon as you have pressed the 'stop' button you are to slide the screen to the left, like this, read aloud the number on the display, and then slide the screen back again [demonstrating]. Now the important thing at this point is that you decide wether the number is odd or even (0,2,4,6,8 are even numbers, and 1,3,5,7,9 odd numbers).

After saying the number, you are to turn up the front card of this set [demonstrating]. Printed on the surface of each of the cards there is another number, likewise between 0 and 9. After reading the number, silently, just place the card here on the table [indicating] with its back up. Then your task will be to decide whether the number on the card was the same or was different from the one you just produced on the counter, that is, with regard to THE ODD/EVEN PROPERTY. For instance, if one number is 5 and the other 9, then they are 'same' with regard to odd/even. But if one is 2 and the other 7, then they are 'different' regarding odd/even. Do you understand?

As soon as you have made this comparison, please score the result on this protocol sheet: make a mark here [indicating] if the numbers were 'same' and here if they were 'different', and then immediately say 'yes'. Then, after 3 seconds, I will again say 'number'. This is the signal for you to repeat the same manoeuvres as before: produce a number on the counter, read it aloud, turn up another card, but this time from this second set [indicating], compare the two numbers with regard to odd/even, mark the result of the comparison ('same' or 'different') in the protocol. Then immediately say 'ready'. And that completes one trial.

Then, after about 15 seconds I will again say 'number', which signals the start of the next trial, in which you are to follow exactly the same procedure as before. Now, do you have any questions?..."

APPENDIX II

Effect of the alternating switch on RNG

The effect of an alternating switching action upon a biased RNG in terms of the probability (P) of one outcome being followed by a different outcome may be calculated:

Pc1 = P of a congruent first outcome
Pc2 = P of a congruent second outcome
Pi1 = P of an incongruent first outcome
Pi2 = P of an incongruent second outcome
Pxy = P of the sequence xy
P = P without an alternating switch
P′ = P with an alternating switch included

The probability under switching conditions (switch on every 2nd trial) for obtaining a sequence of two different conditions then becomes:

$$P'(i1c2,c1i2) = (1/2)(Pc1*Pc2+Pi1*Pi2) + (1/2)(Pc1*Pi2+Pi1*Pc2),$$

where the left sum is P for sequences beginning with a non-switched trial, the right sum P for the sequences beginning with a switched trial. As can be seen, the value of P′ is fairly insensitive to imbalance between Pc and Pi; thus, for example, if we assume Pc=2/3 (Pi=1/3), the values of P′(i1c2,c1i2) increase from .50 to .56 as compared to the case of perfect balance (Pc=Pi=.5).

EUR.J.PARAPSYCHOL. V.5, 1983

PSYCHOKINESIS EXPERIMENTS IN CASEIN INDUCED AMYLOIDOSIS OF THE HAMSTER

Frans Snel P.R. Hol
University of Utrecht

ANIMAL PK MODELS

Only a few PK experiments with disease models of mammals have been conducted. The topics examined - all using mice - were: tumorogenesis (Elguin and Onetto, 1966), woundhealing (Grad, 1961; 1965), resuscitation of anaesthetized animals (Watkins and Watkins, 1971; Wells and Klein, 1972) and psi expectancy effects with malarial mice (Solfvin, 1982).

It is of importance to find out if psi contributes a substantial part of healing, and if so, to find out with which parameters it might be possible to detect that part. There has been little work done on the usefulness of clinical parameters for parapsychology, e.g. measurements in fluids or tissues of living systems which are easily performed, reliable and repeatable. We had the opportunity to collaborate in an experiment using hamsters as a model for the pathogenesis of amyloidosis which allowed us to investigate whether psychic healers could influence this experimental disease in hamsters.

HYPOTHESIS

In these experiments, it was hypothesized that target animals remotely treated by healers would show parameter values which could be interpreted as being 'better' than the values obtained for the controls. The idea was that these parameters could give information about the state of 'well-being' of the animals. If true, these parameters might indicate a useful tool for further research in healing.

THE DISEASE MODEL

Amyloidosis is mainly a disease of older age with unknown aetiology and for which at the moment no treatment exists. It is defined by pathologists as a degenerative change in vertebrates characterized by extracellular deposits of amyloid. Amyloid is histologically a homogeneous and proteinaceous material and does not induce any inflammatory response. Electron-microscopically it is characterized by rigid non-branching fibrils, the amyloid fibrils. Due to the deposition of amyloid in the extracellular spaces, cell function is compromised or fails, resulting in clinical symptons and finally death. At present, several different forms of amyloidosis are recognized (see Gruys et al, 1981, table 1).

In several species amyloidosis occurs as a disease secondary to spontaneous or induced inflammatory lesions (Gruys, 1979a), in which cases it is called secondary or reactive amyloidosis (Glenner, 1980). It is found in man, especially in this 'secondary' form, in patients with tuberculosis, rheumatoid arthritis (Cohen, 1978; 1980; Franklin, 1977; 1980) and paraplegia (Tribe, 1969). The major component in secondary amyloidosis, protein AA, is a unique component showing no homology with immunoglobulins or with any other protein so far identified (van Rijswijk, 1981). It is similar in different patients and is homologous in various species, i.e., man, monkey, mink, mouse, guinea pig, rabbit, duck and cattle (Gruys, 1979a). It shows immunological cross reactivity with a serum alpha-globulin, SAA, the levels of which fluctuate in a manner analogous to acute phase reactants. Relatively large quantities of SAA are formed in the liver, which is the site of origin of many other acute phase reactants (Gruys, 1979a; Fisher and Gill, 1975). Recently, SAA was found to be a

TABLE 1
(after Gruys et al, 1981)
Current chemical classification of amyloid

Type of amyloid	Characteristic protein	Precurser protein	Associating conditions or site
AA	protein AA	SAA	secondary to inflammatory diseases or idiopathic
AL	light chain proteins	para-protein	secondary to plasma cell dyscrasia or idiopathic
AE		(pro)-hormon	APUD-endocrine tissue related amyloid
AS clc2			senile cardiac amyloid: isolated atrial amyloid (IAA) and senile cardiac amyloid (SCA)
AS blb2			senile brain amyloid, vascular and plaques
AD			cutaneous amyloid
AF p	prealbumin-like protein	neural pre-albumin	Portugese familial amyloidotic poly-neuropathy (FAP)
AF j Arao = AFp Ogawa			Japanese familial amyloidotic poly-neuropathy (FAP) in Arao city and Ogawa village

high density lipoprotein (Gruys and Timmermans, 1980; Bendit and Erikson, 1977, 1979) of which the apolipoprotein , apo SAA, contains a protein AA like fragment (Bendit et al, 1982; Hoffmann and Benditt, 1982a; 1982b).

In this paper we are concerned with induced amyloidosis associated with chronical inflammatory processes. Deposits of amyloid have been found in almost every organ in the hamster, e.g. spleen, liver and kidney. Two experiments are described.

THE FIRST EXPERIMENT

Healers

Four natural healers or 'magnetizers' (a term commonly used in Holland to describe a mental or laying-on-of-hands type of healer (Solfvin, 1982)) were contacted by the first author from a list of ten. They were all members of the N.F.P.N., the major Dutch healers association. (note 1)

All healers received photographs of the target cages, which were randomized and distributed as described below. There were seven experimental photographs with five hamsters in each cage; the photographs of the control hamsters were kept at the Parapsychology Laboratory, Utrecht. One healer tried to prevent the onset and severity of the disease during the entire experiment (30 days); three others worked for one week prior to the expected appearance of amyloid (day 16 till day 23). The animals were treated by the healers once a day, at a time that suited them best.

Randomization

A total of 100 hamsters were involved in the study. They were housed in 20 cages holding five hamsters each. Each cage was identified by a letter (A through T). Hamsters within a cage were not distinguished from one another and the experimental and control animals were assigned by cages. Six cages were randomly excluded from the experiment for other purposes. The remaining 70 hamsters were randomly divided (by G.F. Solfvin and B. Millar) into 14 groups of five. The 14 cages were marked with a letter, of which seven were targets and seven controls. All cages were photographed and randomly assigned to the experimental (healing) or control (non-healing) conditions by blindly shuffling twenty index cards with the cage identifying letters, removing the top six and dealing out the remaining 14 into two equal piles, designated H and C.

Each of the two groups of cards was then repeatedly shuffled and the resulting sequence of letters recorded on paper. This was repeated six times for each group, producing two long lists (H and C) of permutations of the letters. The first three letters were then written, alphabetically arranged, on the first row of the sampling schedule, the second three letters from each list on the second row, and so forth, until the lists were exhausted. A copy of the completed sampling schedule was given to the experimenters.

On the first sampling day, one hamster from each of the six cages indicated by the first row of letters was arbitrarily taken from the cage for sacrificing. The experimenters did not know which cages were from the H and C groups. Subsequent rows of the sampling schedule were used for sampling on subsequent sampling days. Because of the primary experiment, i.e. the pathogenesis of amyloidosis, not all 70 hamsters could be studied in the healing situation; since 28 were needed for other purposes we used 42 animals, 21 in the experimental condition and 21 in the control condition.

Hamsters

Young adult hamsters (Mesocricetus auratus), all males, weight 80-100 grams, were given daily injections of 2 ml casein 5% (Hammarsten in 0.3 M $NaHCO_3$, pH 7.5, stored at 4 degree Celsius) subcutaneously, five days a week (Gruys et al, 1979b). The animals were kept in cages of five and received commercial rodent food (Complete Hamster food, Hope Farms, Woerden, Holland) and water ad lib. Night and day hours were not influenced by artificial means.

Experiment

The animals were sacrificed and samples taken according to a sampling schedule every third injection day. The standard procedure was as follows: 1) The same time in the morning blood was put immediately on ice. 2) Body weight, haemoglobin, red and white blood count, lactate dehydrogenase and gamma-glutamyltransferase were measured or calculated later the same day, usually starting within one hour after sampling. 3) The plasma was then frozen at -20 degrees Celsius for later use (total protein, cathepsin-D, electrophoresis and serum amyloid A). The normal values (n.v.) for all parameters were calculated from six hamsters on day 0. All measurements were done blindly.

The haematological measurements (Hb, RBC, WBC) were made on heparinized whole blood, obtained by cardiac puncture after anaesthetizing the hamsters with phenobarbital. The enzymatical measurements (LHD, gamma-GT, Cath-D) in plasma (which was separated by centrifugation - 20 minutes, 4000G - immediately after puncture) and the measurements of total protein, SAA and electrophoresis were performed on frozen stored plasma after thawing. Bloodsmears were obtained from peripheral (toe) blood for the differentiation of white blood cells and counting of the platelets.

Analysis

The effectiveness of healing on each of the eleven clinical parameters was explored by means of a two-way factorial analysis of variance. The factor other than experimental/control was the day of sampling; it was expected that this would contribute a considerable amount of variability to the results, so it was treated as a nuisance variable.

THE SECOND EXPERIMENT

Healers

Five healers participated in this experiment, three professionals and two persons who acted as healers (acquaintances who when asked were enthusiastic and interested in the idea of trying to heal the hamsters in this way). Treatment by absent healing was once a day.

Each healer received two photographs of five hamsters.

Randomization

There were ten photographs of five hamsters, two for each healer. Cages were numbered 1 to 10. The photographs were randomized and distributed to the healers by an otherwise uninvolved person. Every sampling day, one hamster was taken from each cage. When the code was broken after the experiment was finished, it showed that on each sampling day 8 hamsters were designated as experimental animals and two as controls. So after all the sampling was done, one hamster from each cage served as a control. The healers could only have known this in a clairvoyant way. All photographs were coded twice to prevent any possibility of a clue.

Hamsters

In this experiment there were 50 hamsters. The same method for inducing amyloidosis as described above was used, with two exceptions: a) the hamsters got casein injections every day of the week for 50 days, except for day 27 when no injection was given; b) sampling was done every tenth day. The hamsters were young adult males, average weight 123 (sd=19) gram, photographed in groups of five. Environmental conditions were the same as in the former experiment. All measurements were done on sampling days.

Experiment

Experimental parameters were body weight, Hb, gamma-GT and LDH. The methods of measuring the parameters were the same as described in the previous experiment, except for LDH. Instead of using a commercially available testkit we made the buffer ourselves (note 2). Measuring was done with 1 ml of this mix + 50 microliter serum at 340 nm. The same apparatus was used as in the former experiment.

Analysis

It was stated in advance that a matched t-test, two tailed (Siegel, 1956) should be used. Subjects were to get feedback only if the results were in the predicted direction.

Parameters

a) The mean weight of all animals was calculated at the beginning of the experiment and on the day they were anaesthetized.

Heamatological measurements

b) The haemoglobin (Hb) level was determined using the 'Haemoglobin Test Combination', No 124.729, of Boehringer, Boehringer Mannheim GmbH Diagnostics, Germany.
c) Total Red Blood Count (RBC) was performed on 1:50,000 diluted blood samples using a Micro Medic automatic pipettor (Micro Medic Systems, Philadelphia) and a Coulter Counter model 2F of Coulter Counter Electronics Ltd., Dunstable, England (Sonnenwirth, 1980).
d) Total White Blood Count (WBC) on a 1:500 diluted sample, with the same apparatus (Sonnenwirth, 1980).
e) Platelets were counted from the bloodsmear per 1000 red blood cells and multiplied with the total number of the RBC divided by 1000 (Sonnenwirth, 1980).
f) Peripheral bloodsmears were prepared for leucocyte differentiation (Diff) and stained according to May Grunwald-Giemsa.

Enzymatical measurements

g) Lactate Dehydrogenase (LDH) was determined at 25 degrees Celsius using LDH test combination No 124.907 of Boehringer. The activity is expressed in units per liter blood plasma.
h) Gamma Glutamyl Transferase (gamma-GT) was determined at 25 degrees Celsius using test combination No 124.702 of Boehringer. The activity is expressed in units per liter blood plasma.

i) Cathepsin D was determined by a modification (Hol, 1983, in press) of the haemoglobin method of Barrett (1971).
j) The measurement of Total Protein (TP) was done according to the method of Lowry (Lowry et al, 1951).
k) The electroforesis (EF) was done by conventional Polyacrylamide Gel Electroforesis (PAGE) with an LKB 2117 Multiphor (LKB Produkter AB, Bromma, Sweden). Horizontal thin layer gels were used containing 7.5% acrylamide buffered with Tris-glycine, pH 8.9. Plasma samples were done according to LKB's Application Note 306. Scanning of the gels was done with an I.L Boskamp Densitometer 377 at 575 nm (Instrumentation Laboratory, Lexington, Mass.).
l) The SAA levels were measured with an enzyme linked immunosorbent assay (ELISA). AA antigen was coated on DynaTech Microtiter plates. In this assay the competition was measured between the coated AA antigen and the SAA antigen in test serum against specific rabbit anti-AA IgG. The microtiter plates were scanned with a Titertek Multiscan (Flow Laboratories, Irvin Ltd., Ayrshine).
m) The amyloid deposition in liver, spleen and kidney was quantified using a semiquantitative method: from -(negative), +-, ++++ (largest quantities of amyloid). The mean score per group of animals from liver, spleen and kidney taken together are given graphically tabulated from 0 to 5 (figures 1 and 2).

It was stated that for this experiment healthier values would mean:
1) a higher weight, Hb and RBC.
2) lower WBC and platelet counts, in the differentiation especially a lower number of bandforms (meaning a less severe infection), LDH, gamma-GT and Cath-D levels.
3) as for TP and EF, we expected the values to be as close as possible to the normal values.

RESULTS

Clinical observations in the hamsters

During the induction of amyloidosis three periods could be discerned: one in which all animals were negative for amyloid (14 days), a transitional period and a period in which all animals were positive for amyloid (after 21 days). Generally, during the first two weeks heavy inflammatory lesions, abcesses and ulcers on the injection

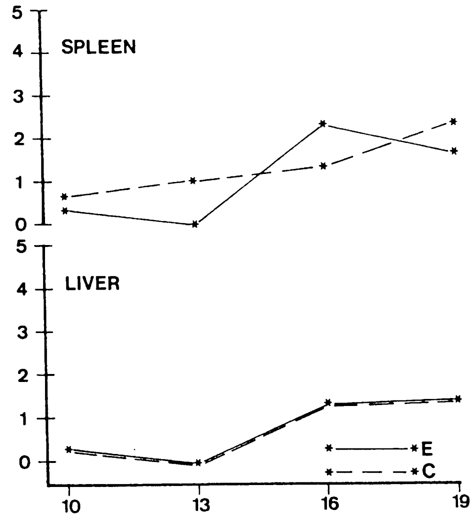

FIGURE 1
Mean score amyloid deposition in liver and spleen
First experiment. Experimental n=3, Control n=3.

sites were obvious. These reached a peak from day 7 till day 14. After two weeks a recovery from these macroscopic lesions appeared, despite further injections. In the first days of the experiment, the WBC showed a rapid increase after which a plateau was formed. The number of platelets and the activity of Cath-D increased as well. The Cath-D activity decreased after two weeks and increased again at the beginning of the forming of amyloid deposits (Hol, 1983, in press). Eosinophilic and basophilic cells in the differentiation of the white blood cells were infrequently seen and are not represented.

Deposition of amyloid was found in liver and spleen, and in the second experiment also in the kidney. The scoring of results is presented in the figures 1 and 2. The values of the SAA measurements are not presented here, as there were only a few samples in some of the experimental groups. Furthermore, in the second experiment each datapoint for the control group was based on two measurements or less, so it was not considered worthwhile to pursue the matter statistically. These results, however, were consistent with patterns

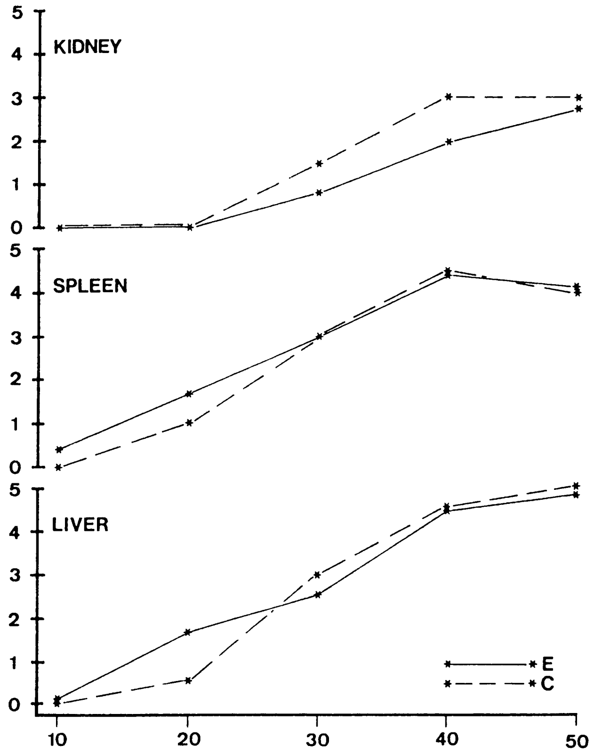

FIGURE 2
Mean score amyloid deposition in liver, spleen and kidney
Second experiment. Experimental n=8, Control n=2.

obtained in other 'non parapsychological' experiments (HOL, to be published).

Healing effects first experiment

The results of the measurements are presented in tables 2 to 12, which indicate the following:

1) Day: days on which the casein injections were given.
2) Experimental and control column: values given are the means, with the standard deviations in parentheses
3) n.v. indicates the clinical normal values for each parameter in the range of mean minus 2 standard deviations to mean plus 2 standard deviations.
4) the bottom lines give the results of the analyses of variance.

TABLE 2
Results of weight measurements, first experiment
(injection days)

Day	Experimental	vs Control group
0	95.8 (1.5)	83.3 (10.5)
4	83.5 (2.6)	87.1 (1.4)
7	83.6 (8.8)	88.7 (5.0)
10	83.8 (9.9)	78.3 (5.1)
13	76.2 (8.0)	82.8 (5.5)
16	83.8 (7.9)	77.0 (5.1)
19	82.8 (4.6)	79.3 (5.1)

normal value: 70 - 109 g

anova: df=1, F=0.213, p=0.648

Seven of the parameters did not show any influence of healing: weight (p=0.65), red blood count (p=0.99), white blood count (p=0.29), platelets (p=0.35), differentiation for the segmented cells (p=0.37), lymphocytes (p=0.21), monocytes (p=0.40), cathepsin-D (p=0.42) and total protein (p=0.42).

TABLE 3
Results of heamoglobin measurements, first experiment
(injection days)

Day	Experimental vs	Control group
0	9.3 (0.3)	9.2 (0.3)
4	9.6 (0.4)	9.0 (0.2)
7	8.2 (0.5)	8.0 (0.4)
10	7.3 (0.1)	7.4 (0.2)
13	7.6 (0.3)	7.6 (0.4)
16	7.5 (0.4)	7.6 (0.2)
19	7.4 (0.2)	7.0 (0.3)

normal value: 8.7 - 9.7 mmol/L

anova: df=1, F=3.936, p=0.057

TABLE 4
Results of red blood count, first experiment
(injection days)

Day	Experimental vs	Control group
0	7.7 (0.4)	7.8 (0.1)
4	8.3 (0.2)	8.6 (0.4)
7	7.6 (0.4)	7.8 (0.5)
10	7.3 (0.4)	7.5 (0.4)
13	7.2 (0.4)	7.0 (0.3)
16	7.3 (0.3)	7.3 (0.4)
19	7.3 (0.1)	6.9 (0.2)

normal value: 7.1 - 8.3 * 10 **12/L

anova: df=1, F=0.000, p=0.986

TABLE 5
Results of white blood count, first experiment
(injection days)

Day	Experimental	vs Control group
0	3.8 (0.4)	5.1 (0.6)
4	13.8 (2.9)	11.3 (0.5)
7	13.8 (1.1)	17.8 (9.2)
10	13.3 (1.0)	16.6 (1.6)
13	8.7 (1.5)	10.4 (3.4)
16	14.8 (5.4)	17.2 (1.7)
19	9.4 (3.3)	12.3 (3.1)

normal value: 2.9 - 6.1 * 10 **9/L

anova: df=1, F=1.144, p=0.294

TABLE 6
Results of platelet counts, first experiment
(injection days)

Day	Experimental	vs Control group
0	131 (11.5)	164 (28.9)
4	199 (33.1)	283 (17.7)
7	178 (19.0)	194 (24.0)
10	304 (49.8)	369 (80.8)
13	613 (99.4)	546 (98.9)
16	560 (79.3)	580 (95.5)
19	504 (37.2)	527 (23.5)

normal value: 92 - 202 * 10 **9 /L

anova: df=1, F=0.920, p=0.346

TABLE 7
Results of counts of leucocyte differentiation, first experiment (injection days)

Day	Bandforms Exp.	vs Contr.	Segmented Exp.	vs Contr.
0	0.0 (0.0)	0.3 (0.5)	21.7 (4.5)	23.3 (4.5)
4	10.0 (0.8)	14.7 (2.6)	51.0 (2.9)	41.7 (0.9)
7	9.7 (2.6)	13.7 (1.0)	37.7 (7.4)	49.0 (8.3)
10	7.7 (1.3)	7.7 (1.3)	35.3(10.7)	33.0 (6.4)
13	4.5 (0.5)	6.8 (1.5)	29.8 (7.6)	40.5 (5.4)
16	7.6 (2.1)	7.6 (1.6)	29.2 (9.7)	33.0 (4.6)
19	6.0 (1.4)	6.0 (2.2)	58.0 (4.3)	60.0(11.1)
normal value:	0 - 1 %		13 - 32 %	
anova: (df=1)	F=6.450, p=0.017		F=0.837, p=0.368	

Day	Lymphocytes Exp.	vs Contr.	Monocytes Exp.	vs Contr.
0	68.7 (4.8)	62.7 (4.6)	8.7 (2.6)	13.0 (1.6)
4	32.7 (0.9)	33.7 (1.2)	6.3 (2.1)	9.7 (2.4)
7	39.3 (8.5)	25.7 (5.2)	13.3 (1.3)	12.0 (3.3)
10	43.7 (6.1)	48.0 (7.5)	12.3 (5.2)	11.3 (2.1)
13	58.0 (5.8)	46.0 (5.5)	7.5 (3.0)	6.5 (3.2)
16	42.8 (7.3)	44.8 (6.1)	19.4 (6.0)	13.4 (3.9)
19	31.3 (4.6)	60.0(11.1)	4.7 (0.9)	2.0 (0.8)
normal value:	55 - 77 %		5 - 17 %	
anova: (df=1)	F=1.668, p=0.207		F=0.727, p=0.401	

TABLE 8
Results of lactate dehydrogenase measurements, first experiment (injection days)

Day	Experimental	vs Control group
0	79.7 (5.4)	59.3 (3.8)
4	149.7 (34.1)	146.0 (44.7)
7	110.0 (24.1)	165.0 (55.8)
10	143.0 (46.3)	218.7 (55.5)
13	60.0 (6.6)	137.3 (66.7)
16	82.8 (23.2)	98.2 (17.7)
19	64.7 (15.5)	57.7 (12.1)

normal value: 47.1 - 91.9 U/L

anova: df=1, F=4.841, p=0.036

TABLE 9
Results of gamma glutamyl transferase measurements, first experiment (injection days)

Day	Experimental	vs Control group
0	2.7 (0.5)	5.0 (0.0)
4	9.7 (1.3)	11.5 (1.5)
7	12.0 (2.9)	13.0 (2.0)
10	9.7 (2.4)	15.0 (3.3)
13	--	--
16	14.2 (8.8)	16.8 (12.0)
19	7.0 (2.2)	9.0 (2.9)

normal value: 1.2 - 5.0 U/L

anova: df=1, F=3.580, p=0.069

TABLE 10
Results of cathepsin - D measurements, first experiment
(injection days)

Day	Experimental	vs Control group
0	21.6 (2.3)	19.0 (2.7)
4	37.9 (2.4)	37.3 (4.1)
7	29.8 (6.8)	33.8 (4.4)
10	29.8 (1.3)	23.4 (3.3)
13	14.7 (1.2)	19.9 (8.1)
16	23.6 (5.8)	19.9 (3.1)
19	27.2 (2.6)	25.7 (4.1)

normal value: 14.7 - 25.9 u/min/ml

anova: df=1, F=0.659, p=0.424

TABLE 11
Results of total protein measurements, first experiment
(injection days)

Day	Experimental	vs Control group
0	39.6 (2.8)	40.9 (2.4)
4	45.8 (0.5)	44.7 (0.7)
7	36.0 (1.8)	42.8 (3.9)
10	45.1 (5.4)	42.1 (1.6)
13	41.0 (1.2)	39.6 (1.7)
16	45.1 (0.9)	44.7 (0.7)
19	43.6 (2.6)	43.9 (1.9)

normal value: 34.9 - 45.5 g/L

anova: df=1, F=0.255, p=0.618

TABLE 12
Results of electrophoresis measurements, first experiment
(injection days)

	alfa			beta		
Day	Exp.	vs	Contr.	Exp.	vs	Contr.
0	7.0 (0.8)		8.0 (0.8)	29.3 (2.6)		30.0 (3.6)
4	6.3 (0.5)		5.3 (1.3)	20.0 (3.6)		14.3 (0.9)
7	7.7 (2.1)		8.3 (3.1)	24.3 (6.1)		24.3 (8.1)
10	7.3 (1.7)		9.0 (2.2)	24.0 (2.2)		26.0 (1.6)
13	6.0 (1.6)		3.5 (2.7)	32.0 (2.6)		34.0 (4.1)
16	9.4 (0.5)		10.0 (0.6)	28.2 (1.9)		27.0 (1.7)
19	8.3 (1.3)		12.3 (3.7)	25.3 (3.4)		27.7 (2.5)
normal value:	5.6 - 9.4 %			23.4 - 36.0 %		
anova: (df=1)	F=1.092, p=0.305			F=0.001, p=0.972		
	gamma			**albumin**		
Day	Exp.	vs	Contr.	Exp.	vs	Contr.
0	23.7 (4.1)		22.7 (5.3)	38.0 (1.4)		37.0 (2.2)
4	22.0 (4.3)		15.3 (0.9)	47.7 (8.1)		63.0 (3.6)
7	21.3 (2.6)		18.3 (3.1)	44.3 (5.7)		45.3(11.8)
10	25.7 (1.3)		26.3 (1.7)	40.0 (2.2)		35.3 (2.5)
13	41.8 (4.7)		41.8 (3.5)	17.5 (3.8)		18.0 (5.2)
16	21.0 (2.3)		20.8 (1.2)	40.4 (2.4)		40.2 (1.2)
19	23.3 (2.5)		22.0 (3.7)	41.7 (1.9)		36.3 (3.3)
normal value:	13.6 - 32.8 %			33.7 - 41.3 %		
anova: (df=1)	F=1.610, p=0.215			F=0.143, p=0.708		

Four parameters showed values of some interest: haemoglobin (p=0.057) and gamma-GT (p=0.069) approach significance; lactate dehydrogenase (p=0.036) and in the differentiation the bandforms (p=0.017) are significant. These last four values were the motivation to cooperate in a second experiment in which these parameters could be studied further.

Healing effects from second experiment

Three parameters from the first experiment were studied in this second one: Hb, LDH and gamma-GT. We did not repeat the differentiation of the white blood cells because we considered the results too dependent on the varying daily interpretations of the analyst. Body weight was added as a parameter to study the possible correlation between the loss of body weight and the progress of the disease.

The results of the combined five sampling days showed no significant effect in any of the parameters (table 13).

TABLE 13
P values from second hamster experiment, between groups

Group	weight	Hb	LDH	-GT
Experimental hamsters vs Controls	>0.20	>0.20	0.20	>0.20
professional healers vs people acting	>0.20	>0.20	<0.20	>0.20
professional healers vs controls	>0.20	>0.20	>0.10	>0.20
acting people vs controls	>0.20	>0.20	>0.20	>0.20

When looking into the results per sampling day, the hypothesis was rejected. There were clinically no interesting improvements in the hamsters' health; to the contrary, on day four the level of haemoglobin was significantly lower in the treated animals (p<.05

two-tailed). As an indication of 'healthier' values this result is not encouraging.

DISCUSSION

The professional healers as well as the two people acting as healers were approached by the first author. They were fully informed about the experimental conditions and procedures well before the experiments began. Some of the healers did not approve of the experimental conditions; that is, it is not part of their philosophy to give animals injections to induce a disease and then try to treat the animals. After some deliberation, however, they agreed to treat the hamsters the best they could. Casein injections produce inflammatory reactions and ulcers, in the first period of 14 days (the predisposition period). The inflammatory parameters white blood count, differentiation and enzyme levels (Cathepsin-D included) change as indicators of cell damage. Less cell damage, as indicated by these parameters, could be an effect of healing. In the disease period, high amyloid scores indicate a disease effect; less amyloid deposition should then be a healing effect. As has been shown in the results no healing effect was found for the inflammatory parameters and amyloid quantities. As shown in tables 3,7,8, and 9, four parameters gave some encouraging results in the predicted direction.

As a consequence we cooperated in a second experiment. The results of this second experiment, however, were very poor. The one significant p value for the Hb value (Hb, $p<0.05$, two tailed) was opposite to the predicted direction. With so many variables we did not expect all the results to be conclusive. However, one finding is consistent with an experiment where the haemoglobin levels in human subjects were measured after treatment by laying on of hands as compared with a control group (Krieger, 1976). She found that when 32 nurses who were trained as healers treated hospitalized patients with a variety of diseases, the haemoglobin levels rose significantly: $p<.001$. These results, however, do not necessarily represent a healing effect in a parapsychological sense. Personal contact and attention as a result of the increased time spend with the sick human being might be considered as the most important factor in a psychological sense.

One could reason that in the second experiment the randomization procedure (designating which were to be the experimental and control animals after the experiment was finished) gave the healers little chance to produce effects. There is also the factor that the injections were given every day, possibly leaving the healers only a slight chance of healing. One other argument to consider is that some of the healers did not approve of the experimental conditions. They could therefore be influenced by negative feelings during the experiment. It could also be that absent healing (by means of a photograph) is not that simple. All healers complained about the lack of opportunity to get 'through' to the animals; they deplored it not being possible to touch or see the animals in their surroundings. Still another possibility is that the strict clinical and experimental conditions prevented a healing effect.

Suggestion does not seem to be a good argument for explaining these results, because:
1) it does not seem easy to suggest 'healthier' values to hamsters by means of absent healing.
2) none of the experimenters thought it possible to influence an induced disease, as distinct from a natural disease, with daily repeated injections.
3) the measurements were done blindly.

Another factor could have been stress through handling and injecting the hamsters every day. These are, however, standardized procedures and were equal for all hamsters. Environmental factors could be thought of as an influencing factor, but the hamsters were kept in the same room and were subject to the same conditions throughout the experiment. In sum, in these experiments the standardized conditions did not facilitate a healing effect. More research with living non-human systems is necessary to study the possible value of clinical parameters for parapsychology.

ABSTRACT

Two experiments with hamsters were conducted in which a number of clinical parameters were studied for possible influences from absent healing by natural healers and people acting as healers. They tried to influence the course of an induced disease (casein induced secondary amyloidosis) with a photograph as inductor. A few suggestive results were obtained in the first experiment which were not confirmed in the second experiment. It is discussed whether absent healing can

influence an induced disease in a living non-human system.

NOTES

1. To become a member of this association one has to pass an examination before a board of the organization and, when accepted, one can start as an apprentice healer. Patients treated by the apprentice healer are interviewed by members of the board and, when satisfactory results are obtained, the healer can become a full member. Members are required to uphold certain standards and are allowed to charge a (low) fee for their work.

2. A 50 mM KPi buffer was made - containing 2.1 g K_2HPO_4 (BDH Chemicals, Poole, England) and 270 mg KH_2PO_4 (Merck, Darmstadt, Germany) in 270 ml aqua dest -, pH 7.5. To 100 ml of this buffer 6.2 mg Natrium pyruvaat (Serva, Heidelberg, Germany) and 15 mg NaDH (Boehringer, Mannheim, Germany) were added. The pH was controlled and, when necessary, adjusted with 1 N HCL.

ACKNOWLEDGEMENT

Many people have assisted us with these laboratory studies. We are especially grateful to Prof. Dr. E. Gruys (Faculty of Veterinary Medicine, Department of Pathology, University of Utrecht) for help and advice during the experiment and to Brian Millar Ph.D. and G.F. Solfvin (M.A.) (both at that time at the Parapsychology Laboratory, University of Utrecht) who did the randomizations and helped with the statistics. In the laboratory, the assistance of Petra Punt, Herman van Rooyen and Anne Marie van Ederen was invaluable.

REFERENCES

Barrett, A.J. 'A new assay for Cathepsin B1 and other Thiol Proteinases', Anal. Biochem., 1972, 47, 280-293.

Benditt, E.P. and Erikson, N. 'Amyloid protein SAA is associated with high density lipoprotein from human serum', Proc. Natl. Acad. Sci. USA, 1977, 74 (9), 4025-4028.

Benditt, E.P. and Erikson, N. 'Amyloid protein SAA is an apoprotein of mouse plasma high density lipoprotein', Proc. Natl. Acad. Sci. USA, 1979, 76, 4092-4096.

Benditt, E.P., Hoffman, J.S., Erikson, N., Parmelee, D.C. and Walsh, K.A. 'SAA, an apoprotein of HDL - its structure and function', Ann. N.Y. Acad. Sci., 1982, 389, 183-189.

Cohen, A.S., Cathcart, E.S. and Skinner, M. 'Amyloidosis - current trends in its investigation', Arthritis and Rheumatism, 1978, 21 (1), 153-160.

Cohen, A.S. and Wegelius, O. 'Classification of amyloid: 1979-1980', Arthritis and Rheumatism, 1980, 23 (6), 644-645.

Elguin, G.M. and Onetto, B. 'Psychokinesis in experimental tumorogenesis', J.o.P., 1966, 30, 220 (abstract).

Fisher, C.L. and Gill, Ch.W. 'Acute Phase Proteins', in Ritzmann, S.E. and Daniels, J.C. (eds) 'Serum protein abnormalities', Boston, Little, Brown and Company, 1975, 335-337.

Franklin, E.C. 'Amyloidose', Fortschr. Med. 1975, 95 (20), 1329-1332.

Franklin, E.C. 'Immunopathology of the amyloid disease', Hosp. Pract. 1980, 15 (9), 70-77.

Glenner, G.G. 'A retrospective and prospective overview of the investigations on amyloid and amyloidosis - the beta fibrilloses', in Glenner, G.G., Pinho e Costa, P. and Falcao de Freitas, A. (eds) 'Amyloid and amyloidosis', Proc. 3rd Int. Symp. on Amyloidosis, Portugal 1979, Excerpta Medica, Amsterdam, 1980, 3-13.

Grad, B. et al 'The influence of an unorthodox method of treatment on woundhealing in mice', Int. J. Parapsychol., 1961, 3, 5-24.

Grad, B. 'Some biological effects of the 'Laying on of hands': a review of experiments with animals and plants', J.A.S.P.R., 1965, 59, 95-127.

Gruys, E. 'Comparative approach to secondary amyloidosis', Dev. Comp. Immunol., 1979a, 3, 23-36.

Gruys, E., Timmermans, H.J.F. and van Ederen, A.M. 'Deposition of amyloid in the liver of hamsters: an enzyme-histochemical and electron-microscopical study', Laboratory Animals, 1979b, 13, 1-9.

Gruys, E. and Timmermans, H.J.F. 'Bovine renal amyloidosis as a model for secondary amyloidosis and idiopathic amyloidosis of the secondary type', 7th ICLAS Symp., Utrecht, 1979, Gustav Fisher Verlag, Stuttgart, New York, 1980, 139-144.

Gruys, E., Syens, R.J. and Biewenga, W.J. 'Dubious effect of dimethyl-sulphoxide (DMSO) therapy on amyloid deposits and amyloidosis', Veterinary Research Commun., 1981, 5, 21-32.

Hoffman, J.S. and Benditt, E.P. 'Changes in high density lipoprotein content following endotoxin administration in the mouse. Formation of serum amyloid protein-rich subfractions', J. Biol. Chem., 1982a, 257 (17), 10510-10517.

Hoffman, J.S. and Benditt, E.P. 'Secretion of serum amyloid protein and assembly of serum amyloid protein-rich high density lipoprotein in primary mouse hepatocyte culture', J. Biol. Chem. 1982b, 257 (17), 10518-10522.

Hol, P.R., van Ederen, A.M., Snel, F.W.J.J. and Gruys, E. 'Activities of lysosomal enzymes in blood plasma of hamsters during induction of amyloidosis', Proceedings of the Fourth International Symposium on Amyloidosis, Bristol, England, 1981, in press, 1983.

Hol, P.R., van Ederen, A.M., Snel, F.W.J.J, Gruys, E., Langeveld, J.P.M, and Veerkamp, J.H. 'Activities of lysomal enzymes and the levels of SAA protein in bloodplasma of hamsters during casein induction of AA amyloidosis', to be published.

Krieger, D. 'Healing by the 'laying on of hands' as a facilitator of bioenergetic change: the response of in-vivo human haemoglobin', Psychoenerg. Syst., 1976, 1, 121-129.

Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. 'Protein measurement with the folin phenol reagent', J. Biol. Chem., 1951, 193, 265.

Rijswijk, M.H. van 'Amyloidosis', Ph.D. Thesis, University of Groningen, 1981, p.13.

Siegel, S. 'Nonparametric statistics for the behavioral sciences', London, McGraw-Hill, 1956.

Solfvin, G.F. 'Psi expectancy effects in psychic healing studies with malarial mice', E.J.P., 1982, 4, 159-197.

Sonnenwirth, A.C. and Garrett, L. 'Gradwohl's Clinical Laboratory Methods and Diagnosis', London, The Mosby Company, 1980, Vol. I.

Tribe, C. 'Amyloidosis in chronic paraplegia', in 'Renal failure in paraplegia', London, Pitman Med. Publ.Ltd., 1969, 54-90.

Watkins, G.M. and Watkins, A.M. 'Possible PK influence on the resuscitation of anaesthetized mice', J.o.P., 1971, 35, 257-272.

Wells, R. and Klein, J. 'A replication of a 'Psychic healing paradigm'', J.o.P., 1972, 36, 144-149.

Frans Snel
Postbus 264
3720 AG Bilthoven
The Netherlands

P.R.Hol
Faculty of Veterinary Medicine
State University of Utrecht
Utrecht
The Netherlands

EUR.J.PARAPSYCHOL., V.5, 1983

NEW INFORMATION FAVORING A PARANORMAL INTERPRETATION IN THE CASE OF RAKESH GAUR

Satwant K. Pasricha
National Institute of Mental Health and Neuro Sciences,
Bangalore (India)

Rakesh Gaur is the subject of a case of the reincarnation type investigated and reported by D.R. Barker and myself (Pasricha and Barker, 1981). In our separate discussions of this case Barker and I agreed concerning the main facts to be interpreted, but differed on the matter of wether Rakesh had made a sufficient number of statements (and sufficiently precise ones) about the person whose life he claimed to remember before he had been taken to the town (Tonk) where that person had lived.

In the two years following the conclusion of the joint investigation of the case by Barker and myself, I visited again (during 1979-80) many of our previous informants and also interviewed some new ones. But before I discuss my present findings, I think it will be helpful if I summarize here the lines of arguments we followed in our earlier discussion.

In his discussion of the case Barker suggested that Rakesh, before he had been taken to Tonk, had made only "a few general statements about a 'previous life'". For example, Rakesh had said that he had been a carpenter and had died of electrocution. Since he had also mentioned the name Tonk, he was taken there, and a search for a person corresponding to his statements was initiated. Eventually, someone thought of a carpenter called Bithal Das whose life and death (by electrocution) seemed to correspond satisfactorily with Rakesh's statements. Without much reflection on the matter, according to Barker, the informants for both families concerned then exchanged information about what Rakesh had said and about the life of Bithal Das. The result was that much information about Bithal Das was overheard by Rakesh or passed on to him; he then assimilated all this information among his apparent memories of a previous life and later repeated it. The strong conviction of all the persons concerned in the case that Bithal Das was the person about whom Rakesh had been speaking would have tended to reinforce Rakesh's belief that he was Bithal Das reborn, and thus a circular enhancement of a false belief would have occurred.

However, Barker could only make his reference to "a few general statements" by rejecting testimony according to which Rakesh had stated the name of Bithal Das before he had been taken to Tonk. Three major informants, including Rakesh's father, said that he had mentioned this name before going to Tonk, but three others including Bhanwar Lal, Bithal Das's son (Pasricha and Barker, 1981) said that he had not.

In my discussion of the case, I argued that the failure of a child like Rakesh to mention a detail in the presence of one informant should not be used to nullify a statement (from an otherwise reliable informant) that he had mentioned the name on another occasion. Experienced investigators of these cases know that the children subjects mention different details about the previous life to different informants, and they also know that different informants later remember different details about statements they have heard. In my analysis of these cases I do not allow a statement of one informant to cancel that of another (with which it seems to be discrepant) unless they disagree about an event of which they were simultaneous witnesses or unless I find that one informant is, from other evidence, unreliable as a witness.

The interpretation of the case offered by Barker required one or both of two additional assumptions. Rakesh was credited by Bithal Das's son, Bhanwar Lal, with accurate knowledge about numerous details of the home and private life of Bithal Das, which he showed two days after his first visit to Tonk, when Bhanwar Lal came to visit him in Kankroli (the town where he lived). Rakesh also made some mistakes, but these were fewer than the details about which he was correct. Also these fewer mistakes were minor ones and some, like his use of "Arun" (note 1), may not have been mistakes but simply misunderstandings. A normal explanation for the attribution of so much correct knowledge to Rakesh requires that we believe either (a) that Rakesh learned all these details during the single, rather brief meeting between the families in Tonk and that he then rapidly began to personate Bithal Das, or (b) that Bhanwar Lal (when interviewed just a few weeks later by Barker and myself) had grossly distorted what Rakesh had told him. There is no evidence suggesting that either of these assumptions is correct; but, although they seem improbable, neithe can be ruled out as impossible.

Further investigations of the case

During the later interviews (conducted in 1979-80) I obtained some additional information about Rakesh's statements and behaviour related to the previous life. Although some of this information tended to strengthen my belief that Rakesh had paranormal knowledge about the life of Bithal Das, none of it was completely free from the objections, raised by Barker, that Rakesh might have learned normally details of the life of Bithal Das from Bithal Das's own family. For example, according to Bithal Das's son-in-law, Satyanarain, Rakesh had told him that he (as Bithal Das) had given a certain kettle as a gift to Satyanarain's father. Bithal Das had in fact given a kettle to his daughter's father-in-law, Satyanarain's father. This information was certainly not readily available to Rakesh. Therefore, to suppose that he did not have paranormal knowledge about the kettle one must imagine that a member of Bithal Das's family had unwittingly mentioned the kettle and its history to Rakesh before he made this remark.

My later investigations were not, however, mainly intended to obtain information about new details, such as the detail of the kettle that was given as a gift. Instead, I tried to learn wether the informants I interviewed during these later inquiries could still recall details of what Rakesh had said before the two families had met, and wether they could remember what information had passed between the two families at the time of their first meeting, when Rakesh was brought to Tonk in October 1976. In this I was not completely successful, mainly, I think, because of the informants' memories about some details had become somewhat dimmed by the passage of time, as some of them freely acknowledged. Rakesh's mother, for example, could no longer remember wether Rakesh had mentioned the name of Bithal Das before he had first gone to Tonk. And another informant, H.P.Sharma, who had earlier said that Rakesh mentioned the name of Bithal Das (when he had first come to Tonk), in later interviews reversed himself and said that he was not sure that Rakesh had done so.

Under these circumstances I began to ask myself how the case would stand if, in fact, Rakesh had not mentioned the name of Bithal Das before going to Tonk. (Perhaps I should make it clear here that just because I am asking this question does not mean that I have changed my stance and concluded that Rakesh did not mention the name). This raised the question of how one decides that a child subject of one of these cases is referring to only one deceased person and no other. If he mentions enough proper names of people and places, it is usually not difficult to show that, however the child may have obtained the information, what he says can refer to only one person. But if the child makes only "a few general statements" can we from these, considered together, reach reasonable certainty about the correct identification of a person to whose life alone the statements apply?

Indika Guneratne, the subject of a case in Sri Lanka, made a number of rather detailed statements about a previous life in a distant town, but mentioned only two proper names, the name of the town (Matara) and the name of a servant of the person whose life he seemed to be remembering (Stevenson, 1977). These names alone would not have sufficed to identify the person about whom Indika had been talking, but Stevenson argued that the names, combined with other statements that Indika had made, narrowed the possibilities so much that Stevenson believed he had found the single person who alone fitted Indika's statements. Indika's statement that he had owned elephants in the previous life proved of particular importance in solving the case. Stevenson obtained a list of all owners of elephants in the area of Matara and, by working through this list with reference to Indika's other statements, he eliminated all but one of the persons who figured on the list.

It seemed to me it might be possible to approach the analysis of Rakesh's case along similar lines. I decided to consider the case with the assumption that Rakesh had not given the name Bithal Das before he first went to Tonk. Here I may remind my readers that all informants agreed that Rakesh and his family had had no direct contact with Bithal Das's family prior to Rakesh's first visit to Tonk in October, 1976. There had been a slight chance for Rakesh to learn normally about Bithal Das's family when, in the summer of 1976 he briefly met the bus driver from Tonk, Chhittarji; but Rakesh had made his main statements about the previous life before that meeting and on the basis of these statements, and his conversation with Rakesh, Chhittarji immediately thought of Bithal Das. On his return back from Tonk, Chhittarji informed Bithal Das's family about his meeting with Rakesh (the details of which we communicated in our earlier report). The important question was therefore: even if Rakesh did not mention the name Bithal Das, did his statements nevertheless specify one particular person?

How the case would stand if Rakesh had not stated Bithal Das's name

As communicated in our earlier report of the case, all the informants had agreed that Rakesh had made three statements about the previous life before he had gone to Tonk. These were that in the previous life (a) he had lived in Tonk, (b) he had been a carpenter, and (c) he had been electrocuted. These statements were all correct for Bithal Das. The pertinent question is therefore: To what extent can we say that they were true only of Bithal Das and not of anyone else?

The question we have to ask first in this connection is: How frequent were deaths from electrocution in Tonk during, say the fourteen years between 1955 (the year of Bithal Das's death) and 1969 (the year of Rakesh's birth)? Dr. Ian Stevenson (1979) wrote a letter, asking this question, to the Chief Medical Health Officer of Tonk. In response, he recieved a reply (dated April 28, 1979) from Dr.M.M.Gogna, who first explained that he had been Chief Medical and Health Officer in Tonk for only three years during which time there had been two deaths from electrocution. He had, however, searched records farther back and stated that "during the past 8-9 years only 6-7 cases (of death by electrocution) are on record". Presumably, the impression of "6-7 cases" in "8-9 years" arose from doubts about actual cause of death in one case. For example, if a man recieved an electric shock while on a pole and fell off the pole, his death might have been attributed either to the electric shock or to the injuries sustained when he fell down. It would not always be easy to decide between these possible causes of death.

From Dr. Gogna's figures, let us say that seven persons were killed by electrocution in Tonk in eight years. This can be translated into a rate of .88 persons each year or, in slightly rounded figures, twelve persons during the fourteen years that we are considering.

We next have to estimate the number of deaths in Tonk during the fourteen-year period between 1955 and 1969. According to the census for 1970 the population of Tonk was then 43,410 and the reported death rate 2.7 (Vital statistics for India, 1970). We can accept the figure for the population of Tonk as reasonably accurate, but the given death rate appears absurdly low. For the same year, the average death rate for all fourteen towns of Rajasthan with a population of over 30,000 was 9.4; and that for Jaipur, the capital of the state was 14.2. Jaipur is a much larger city than Tonk. It has a medical school and much better hospitals and other facilities for health care than Tonk. Given uniform and accurate registration of deaths, we should expect that the death rate in Jaipur would be appreciably lower, and certainly not higher, than that of Tonk. The low death rate for Tonk in 1970 compared with the rates for Jaipur and most other towns of Rajasthan is almost certainly due to low registration of deaths. Defective reporting was studied by the Registrar General of India and acknowledged in his report (Vital Statistics of India, 1970, p.13). Under-reporting must have been even more marked during the 1950s and 1960s than it was later. If we assume a death rate of 10 per 1,000 for Tonk for the period in question, we shall certainly err in underestimating it. If we now take 45,000 as the average population of Tonk during the fourteen year period we are considering, we shall estimate, with reasonable accuracy, that during this period 6,300 persons died in Tonk.

During the same years (using the rate derived from Dr. Gogna's data) there would have been twelve deaths from electrocutions. But most of these deaths occurred, according to Dr. Gogna, among employees of the State Electricity Board (the publicly owned electricity company) who were working on the electric poles. The considerable hazards for employees of the Electricity Boards in India are well known, and Rakesh's father had assumed that the person about whom Rakesh was talking had been such an employee. This conjecture led him to make the enquiries by mail (note 2) that he addressed to the Electricity Board at Tonk, as mentioned in our earlier report. Rakesh himself, however, had never said that he was an employee of the Electricity Board; on the contrary, he said that he was a carpenter. I think we can reasonably assume that (at the most) three of the twelve men estimated to have been electrocuted in Tonk between 1955 and 1969 were of some trade other than that of professional electricians employed by the Electricity Board. From this we can conclude that among the 6,300 deaths occurring in Tonk during the fourteen year period between 1955 and 1969, three were due to electrocution of persons not employed by the Electricity Board. The chances are therefore approximately 1 in 2,100 that Rakesh's statements referred to someone other than Bithal Das.

However, before accepting the above estimate, we should consider the correctness of two other assumptions, additional to those already mentioned. Were electrocutions more common in the 1950s than they were in the period (1970-79) surveyed by Dr. Gogna when he examined the records of deaths in his department? And would there have been a greater tendency to underreport death from electrocution as a cause of death? I think the first question can be answered negatively. Electrical services have spread appreciably throughout India during the past twenty years. This means that more men have been working on electric poles and exposed to the danger of electrocution. No doubt safety precautions against accidents have improved also, but probably the death rate from electrocution has increased. As for underreporting of such deaths, Dr. Gogna stated in his letter that "most of such deaths are usually reported to the responsible officers". Even if we made a further adjustment for underreported deaths from electrocution, we should still, I think, be entitled to conclude that there is a probability of less than 1 in 1,500 that Rakesh's statements could have correctly applied to someone other than Bithal Das. Therefore, for practical purposes we can say that Rakesh showed paranormal knowledge about Bithal Das before his first visit to Tonk, and it does not make much difference wether he did or did not mention the name of Bithal Das before he went there.

CONCLUDING REMARKS

It is not the main purpose in this communication to insist on my interpretation of the case of Rakesh. I am well aware that the estimates and calculations I have given above require certain assumptions, which I acknowledged as I made them. Behind them lie still other assumptions, such as that one can accept as reliable what all informants agreed Rakesh had said before he first went to Tonk. Even though Barker accepted their statements about three details, more skeptical critics might say that these informants also had later misremembered what Rakesh had said. And even if before Rakesh had gone to Tonk someone had made a written record of what he had said about the previous life that had been preserved - Rakesh's father said he had done this in the postcard he send to the Electricity Board at Tonk - skeptics could still say that the written record had not really been made on the date claimed. In sum, there is no limit to skepticism about these cases - or anything else.

However, one purpose of scientific communications, such as this one, is to enlarge the areas of agreement among reasonable persons. From the base of such an area of agreement further advances can be made. Perhaps then an important aspect of this supplementary note to the case of Rakesh may be the encouragement that I hope it gives to all students of spontaneous cases to ask themselves by what criteria they accept or reject the evidence provided by the informants. Even if all observers of these cases are not yet ready to agree on the same criteria, to state their own criteria as clearly as possible may be a useful exercise, and perhaps even an obligation, for everyone interested in such cases.

NOTES

1. The name 'Arun' was given as his name before by Rakesh in response to Bhanwar Lal's question, "What was your name before?" Actually 'Arun' was a name that Rakesh himself had been given when he was young. Apparently Rakesh seemed to have misunderstood Bhanwar Lal's meaning when he used the word 'before' (meaning before this life, not before he was given his present name, Rakesh).

2. Unfortunately the post card that Rakesh's father had sent, was not available to us at the Electricity Board, Tonk. However, the employees acknowledged that they had received and read the post card, but since the person in question was not an employee in their department, the card was misplaced and eventually had been lost.

ACKNOWLEDGMENT

I wish to express my thanks to Prof. Ian Stevenson for assistance in collecting further information on this case during 1979-80 and his critical comments in the preparation of this paper; to Prof. G.G.Prabhu for his valuable comments; to Prof. H.N.Murthy, Dr. T.N.E.Greville, Mr. V.G.Kaliaperumal and to Mr. D.K.Subbakrishna for statistical advice, and to Ms. Emily W.Cook for helpful suggestions for the improvement of the paper.

REFERENCES

Pasricha,S.K. and Barker,D.R. 'A case of the reincarnation type in India: The Case of Rakesh Gaur', E.J.P., 1981,381-408.

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

Stevenson, I. (1979). Personal communication.

Vital Statistics of India, 1970. Registrar General of India (Vital Statistics Division), Ministry of Home affairs, New Delhi, 1973.

Satwant Pasricha
Department of Clinical Psychology
National Institute of Mental Health & Neurosciences
Bangalore 560029
India

LETTERS TO THE EDITOR

From Carlos S. Alvarado and Alfonso Martinez-Taboas:

In a recent paper on fire immunity in this journal Giovanni Iannuzzo (1982) wrote:

> "All these reports suggest that suggestion, increased motivation or decreased resistance in the subject who is supposed to receive the 'fire-immunity power' should be considered. It does not seem that the paranormal hypothesis suggested by Alvarado (1980) can be considered. A hypothetical PK action by the medium (or the man believing in the phenomena) on the target (the fire) seems improbable. It might reasonably be argued, however, that a possible psi-conductive effect could exist, but the available observations are too insufficient and imprecise to justify such an explanation." (p 273)

I would like to make some comments on Iannuzzo's criticism of my paper (Alvarado, 1980), as well as some general points related to his arguments.

First, it should be clear that I was not referring to fire-walking in my paper (which I decided not to include in my discussion), though the concepts I mentioned to explain the "transference" phenomena may be applied to it.

Second, I did not give special emphasis in my paper to the PK effect criticized by Iannuzzo, which was one of other possible explanations for the "transference" phenomena briefly mentioned. Iannuzzo discusses the issue as if I had supported such idea. I think he also misrepresents my paper when he does not mention that I listed as possible explanations the variables he says should be considered to explain the 'transference' of fire immunity powers: suggestion, increased motivation or decreased (ownership) resistance.

Third, we are not told why a PK action "on the target (the fire) seems improbable", but are asked to accept Iannuzzo's statement with no reasons at all offered to support it. Is the idea considered unparsimonious? Is the area to be influenced (a pit, etc.) considered to be too big or difficult to affect by PK action? Perhaps these may be some of Iannuzzo's reasons.

It may be of interest to notice that F.W.H. Myers speculated on some sort of paranormal control at a molecular level around the heat sources to explain fire immunity (General meeting, 1899, p. 149; Myers, 1903, Vol 2, p 533). Sudre (1956/1960) presents a similar argument. Could thermic action or a thermal source be considered a high lability-low structure target system conceptualized by Braud (1981) to be more susceptible to PK effects than low lability-highly structured targets? I am not trying to defend the PK explanation, but to discuss some relevant theoretical aspects to the concept.

Fourth, Iannuzzo reaches some conclusions regarding the importance of factors involved in explaining fire-walking without mentioning many other variables discussed in the literature he quotes (for a discussion of several non-paranormal hypotheses offered to explain fire-walking see Vesme, 1907, 1928). For instance, no mention is made of the 'spheroidal state' explanation, where the skin's moisture creates a protective layer of vapor at contact with very hot surfaces (e.g., Coe, 1957; General meeting, 1899, p 148). Price (1936) and Brown's (1938) studies are quoted but it is not mentioned that they considered, for example, that the amount of time the sole of the foot was in contact with the heat source was an important variable.

I think that Iannuzzo's paper is useful in pointing out relevant bibliography on the topic under discussion. It is to be hoped that a future paper may explore in more detail the normal and paranormal hypotheses offered to explain apparent fire immunity.

REFERENCES

Alvarado, C.S. 'On the "transference" of psychic abilities: A historical note', E.J.P., 1980, 3, 209-211.

Braud, W. 'Liability and inertia in psychic functioning', in B.Shapin and L. Coly (eds.), 'Concepts and Theories of Parapsychology', New York: Parapsychology Foundation, 1981.

Brown, S.I. 'Three experimental fire-walks', Bulletin IV, University of London Council for Psychical Investigation, 1938.

Coe, M.R. 'Fire-walking and related behaviours', Psychological Record, 1957, 7, 101-110.

'General meeting', J.S.P.R, 1899, 9, 146-150.

Iannuzzo, G. 'Fire immunity and fire-walks: Some historical and anthropological notes', E.J.P, 1982, 4, 271-275.

Myers, F.W.H. 'Human personality and its survival of bodily death', (2 vols.), London: Longmans, Green, 1903.

Price, H. 'A report on two experimental fire-walks', Bulletin II, University of London Council for Psychical Investigation, 1936.

Sudre, R. 'Treatise on Parapsychology', New York: Citadel Press, 1960. (First published in French in 1956)

Vesme, C. de, 'Ordeals', Annals of Psychical Science, 1907, 5, 231-299.

Vesme, C de, 'Histoire du spiritualisme Experimental', Paris: Jean Meyer, 1928.

Carlos S. Alvarado
Division of Parapsychology
Box 152, Medical Center
University of Virginia
Charlottesville, VA 22908
U.S.A.

Alfonso Martinez-Taboas
17 Street 1088
Villa Nevarez Dev.
Rio Piedras
Puerto Rica 00927

Reply from Giovanni Iannuzzo:

I would like to reply to comments of C.S.Alvarado and A.Martinez-Taboas on my paper 'Fire-immunity and fire-walks: Some historical and anthropological notes' (Iannuzzo, 1982).

First of all, I must emphasize that in the above-mentioned work I did not mean to criticize Alvarado's paper on the 'transference of psychic abilities' (Alvarado, 1980), rather the 'paranormal hypothesis' of a PK action by the medium on the fire during fire-walking experiences.

1. In their comments Alvarado and Martinez-Taboas write:

   "...it should be clear that I was not referring to fire-walking in my paper (which I decided not to include in my discussion), though the concept I mentioned to explain the 'transference' phenomena may be applied to it ".

In my paper I did not write that Alvarado was referring to fire-walks; but it should be clear that fire-immunity is one of the most important aspects of fire-walks. This ability is shown by some subjects; it seems that they could transfer their ability to others and, consequently, that it seems to exist a 'transference' of a claimed psi ability. Because in his paper Alvarado discussed the transference of psi abilities, and reported some mediumistic experiences of fire-immunity (Crooks, 1874; Dunraven, 1924; Home, 1921), I took this remark of Alvarado as a starting point for discussing the problem of the nature of fire-immunity. I must, therefore, express surprise for this Alvarado's (and Martinez-Taboas's) comment.

2. I think that in my paper there was no misinterpretation of Alvarado's note. Alvarado listed some possible explanations of transference and among them the PK effect. On the other hand, Alvarado did not express in his paper any criticisms of the PK hypothesis, but merely wrote that this hypothesis also "should be consisdered". I wrote that it does not seem that the paranormal hypothesis (suggested among other hypotheses) by Alvarado can be considered. Where is misinterpretation ?

3. With regards to possible explanations for the transference of fire-immunity, I would like to point out that my statement that "a PK action on the target (the fire) seems improbable" is founded on data reported in the available literature on this topic.

I don't know if Alvarado places some weight on the reliability of all the available literature on fire-immunity. But the primary task here is the preliminary one of deciding if a large number of reported fire-immunity's phenomena should be really trustworthy. I have the impression that this is not true. And, on this subject, I would like to make some comments.

a. Many accounts and relations about fire-immunity's phenomena (and fire-walks) go back to the last decades of XIX century, and the beginning of XX century (see for instance Thompson, 1894 or Lang, 1901-1902), an historical period in which phenomena's controls and investigations were often insufficient. Actually we must properly doubt the truth and reliability of these accounts. Till now no research has been carried out to confirm - or disconfirm - the fire-immunity's claims discussed in the above mentioned literature.

b. Many authors of accounts on fire-immunity and fire-walks were not experienced parapsychologists or physical researchers, but anthropologists, ethnologists, explorers. These authors had reasonably no experience of frauds and, in their writings they seem to be interested fundamentally in describing cultural aspects of ceremonies which they attended. I think that some authors (see, for instance, De Martino, 1949; Sayce, 1933; Ocken, 1894 and others which I prefer not to quote in my discussion) devoted not much attention to the analysis of possible paranormal factors involved in fire-immunity. Are these writings reliable from the parapsychology viewpoint ?

The question is neither a new nor an original one, but when a parapsychologist is faced with a claimed paranormal fact discussed in a paper by another scientist, he must doubt the reliability of it, as well as a psychologist should doubt the reliability of an analysis of a psychological fact carried out by, i.e., a chemist. These remarks are in no way intended to disparage psi researches carried out by others than parapsychologists, but I think it is one thing to quote in writing a review anthropological papers of parapsychological interest and it is quite another to consider them as reliable from the parapsychology point of view. Discipline and restraint are necessary in the case of suspect unsupported paranormal facts, and research on fire-immunity's phenomena have in many instances suffered from a lack of caution and observations.

c. I think that no theories or speculations about fire-immunity in fire-walks today may be considered, for instance speculations of Myers and Sudre on "some sort of paranormal control at a molecular level around heat sources" quoted by Alvarado. In Myers's - or Sudre's - same way, I can hypothize that the fire-walkers's feet should be sheltered from the heat by means of a stuff chemically and/or biologically undetectable (and a similar hypothesis was really suggested by Salvadori, 1946), but my speculation as well as Myers's hypothesis is unsupported by facts. (Myers, 1903; Sudre, 1966).

I would like to repeat that - since now - available data are insufficient to justify a paranormal explanation. For instance, Cassoli (1958) observed the phenomenon but in his work there is the lacking of many important details. He did not record the temperature in the middle or near the edges of the furnace and his observations of the sole of the feet of the fire-walkers after they passed through the furnace seem to be insufficient.
Therefore I think that the idea of PK action on the target (the fire) is unparsimonious, as Alvarado has argued.

4. With regard to the fourth comment of Alvarado (and Martinez-Taboas)

I must emphasize that my paper was a 'note' on historical and antropological aspects of fire-immunity and fire-walks. It should be clear that in a note it is impossible to review carefully all papers, works and variables discussed in the literature, and I decided not to include in my discussion a number of aspects of fire-immunity and fire-walks (for instance, variables discussed in the literature, normal and paranormal hypotheses, psychophysiological aspects). In my opinion the fourth comment of Alvarado and Martinez-Taboas is unjustified: this criticism should be appropriate if in a review (and not in a historical and antropological note) on the topic, I should reach conclusions about the phenomenon without mentioning all variables discussed in literature.

However, I wish to thank Carlos Alvarado and Martinez-Taboas for their comments on my paper. The detailed study of the various aspects of the phenomenon may suggest some theoretical implications for parapsychology, and especially may contribute, in a relevant way to the understanding of the nature of fire-immunity.

REFERENCES

Alvarado, C.S. 'On the "tranference" of psychic abilities: A historical note', E.J.P., 1980, 3, 209-211.

Cassoli, P. 'La pirobazia in Grecia (Le Anastenaria)', Minerva Mediaca, XLIX, 1958, 77, 3679-3686.

Crooks, W. 'Researches in the phenomena of spiritualism', London, James Burns, 1874.

De Martino, E. 'Il mondo magico', Torino, Boringhieri, 1949.

Dunraven, Earl of, 'Experiences in spiritualism by D.D. Home', Proc. S.P.R., 1924, 35, 1-285.

Home, Mme D. 'D.D.Home: His life and mission', London, Kegan Paul

Trench Trubner, 1921.

Iannuzzo, G. 'fire-immunity and fire-walks: Some historical and antropological notes', E.J.P., 1982, 4, 271-275.

Lang, A. 'The fire-walk'. Proc. S.P.R., 1901-1902, 15, 2-15.

Myers, F.W.H. 'Human personality and its survival of bodily death', (2 vols.), London, Longmans, Green, 1903.

Ocken, T.M. 'An account of the Fuji fire-ceremony', Transaction of the New Zealand Institute, 31, 1894.

Salvadori, R. 'Le passeggiate sul fuoco', Milano, Bocca, 1946.

Sayce, R.V. 'Fire-walking ceremony in Natal', Man, 1933, 2, 33.

Sudre, R. 'Trattato di parapsicologia', Roma, Astolabio, 1966 (First published in French in 1956).

Giovanni Iannuzzo
Via S.Salvatore, 9
90018 Termini Imerese (PA)
Italy

THE KOESTLER CHAIR, A LANDMARK IN THE DEVELOPMENT OF OUR FIELD?

The official press release about the establishment of the Koestler Chair at the University of Edinburgh, reads as follows:

The Principal of the University of Edinburgh, Dr. John Burnett, made the following statement this afternoon (Wednesday 22 February 1984):

Koestler Chair of Parapsychology

"Mr. C.E. Robbins, one of the Koestler Trustees, has informed me of their decision to offer to endow a Koestler Chair of Parapsychology, primarily as a research chair, in this University. Personally I am both delighted and excited by this decision and I know that my gratification will be widely shared throughout the University.

The University did not lightly agree when invited by the Trustees to indicate whether or not it would wish to be considered as a recipient of this endowment. However, widespread consultation, undertaken in the light of the controversial nature of parapsychological phenomena, made it clear that there was overwhelming support for the notion both at Senatus and the University Court. In coming to this view, the University noted inter alia that research into parapsychological phenomena was already in progress in the University's distinguished Department of Psychology. This proposal, therefore, provided an opportunity to further objective, scientific research into, and I quote from the Trustees' letter, "the capacity attributed to some individuals to interact with their external environment by means other than the recognized sensory or motor channels" in a critical and well-balanced intellectual climate.

With the clear recognition that such investigations were desirable, intellectually demanding and, whatever their outcome, likely to lead to new understanding, the University indicated its interest. I can only say, on behalf of the University, how grateful I am that the Trustees have recognized the integrity and objectivity of our approach.

The University will discuss details of the endowment, which are not yet known precisely, and subsequent procedures once a formal offer has been received and accepted, both of which are expected to be concluded within a few days"

All actively engaged in our field of inquiry, and here I include a substantial number of our more competent critics, knew that the late, eloquent protagonist of parapsychology, Arthur Koestler, had donated a substantial sum of money for the establishment of a university connected Chair in parapsychology.

The will certainly placed several of the universities within the U.K. in a great dilemma: they potentially wanted the money - especially considering the present situation of a severe financial squeeze - but it also implied officially endorsing the claims of parapsychology, which in Koestler's version were not modest ones. If it just had been a question of accepting that parapsychological experiments be carried out or providing for a lecturer in parapsychology, they would not have hesistated at all. However, Koestler's will made it hard to accept the money without also expressing the degree of recognition that is necessary for the inception of a Chair. At a minimum that implies recognition that parapsychology should be considered as a legitimate area of scientific inquiry, even in the case that the phenomena under investigation turn out to be artifacts of the limitation of present-day scientific methods and thinking, or in the more epistemologicaly revolutionary case, that the phenomena are there but are not in principle easy to manipulate.

The establishment of the Koestler Chair is certainly an important event in our field. The more sanguine protagonists of parapsychology will claim that parapsychology once and for all has made its case in the U.K.. Here we meet that same attitude as the great enthusiasts expressed in 1969, when they learned that the Parapsychological Association had obtained affiliate membership in the A.A.A.S..

The dogmatic critics will hardly see it that way. They will first of all try to discredit the University of Edinburgh, secondly they will try to play down the importance of the event, and thirdly they will try to infiltrate and influence, by political manipulations, the emphasis and content of the work that will be carried out at the established research unit. In the case of most in-group members, by which I mean competent and serious investigators of ostensibly paranormal phenomena (and here I feel that some of the leading and more able critics like Dr. Hyman and Dr. Truzzi could also be included) I conjecture that the news that the University of Edinburgh had accepted the Chair was well received. The University of Edinburgh is by any standard a prestigeous and respected one. In addition, the parapsychological research which has been carried out there by Dr. John Beloff, or under his supervision, enjoys an outstanding reputation for having been carefully and competently executed, including the interpretations of results. This is by and large also the opinion of the critics. Admittedly, these results have been rather palatable to the critics, since the majority have not rendered much evidence of paranormal influence.

The first important step in the establishment of the Chair has been taken and can be considered as a victory for well-balanced parapsychology. It is also a victory and a recognition for the man that pioneered parapsychology in the Department of Psychology at the University of Edinburgh. Already in his book 'The Existence of Mind', published in 1962, he was an advocate for the legitimacy of experimental parapsychology. The main topic of his book was a challange of Gilbert Ryle's opinion that the concept of mind boils down to a linguistic trap - a 'category mistake'. In his book, Beloff expressed the opinion that parapsychology (especially studies in the area of psychokinesis) may turn out to be the ultimate battleground for the mind-body issue. He is well known for his strong position as regards the mind-body issue, and he has adamantly insisted that paranormal phenomena, if they exist, have a bearing on that time-honoured problem. His position is made clear in a most lucid way in his response to a headline published in the THES of February 17th.

"The point is that at the present time, the tacit assumption of experimental psychology and brain science is a physicalistic one. By this I mean the view that everything we do or experience is the result of brain processes which, in turn, are a function of the electro-chemical state of our brain at a certain moment. This view may well turn out to be true but it would be the height of presumption to claim that it must be true. After all, most of us usually think of ourselves as controlling our body rather than the other way around. Certainly if that is true then it is most unlikely that there are any genuinely parapsychological phenomena but, conversely if there are any such phenomena it is unlikely that physicalism is true. Either way we ought to find out. The establishment of the Koestler Chair of Parapsychology in the Department of Psychology at this University is a step in the right direction."

Even if I do not feel very convinced that the small anomalous effects that we sometimes observe, and for lack of a better alternative are subsuming under the heading psi, really have the wide ranging implications that Beloff thinks, critically and competently pursued research into this border area of human knowledge seems to me not only a legitimate endeavour but a 'must'.

The reception of the news about the establishment of the Chair has been mixed. One quite disappointing reaction was the one by John Maddox: No patience for the paranormal (see Nature, March 22, 1984). He may be right in his criticism of extravagant claims made by some of the rather uncritical, strong believers, but his own attitude against the need for further inquiry stands dangerously close to the a priori argument that such phenomena simply cannot exist. One might have the right to expect a more open-minded and sophisticated epistemological position from the editor of the prestigeous publication that he represents.

The next important step to be taken regarding the Koestler Chair will be the appointment of the chairholder. The objectives of the chair should be agreed upon, and it goes without saying that these objectives must not be allowed to be too much in conflict with the intentions that Arthur Koestler expressed. The formation of the committee for the appointment of the chairholder is partly a matter of bringing people together who are considered to have a certain competence for their task. But it is also a policy-matter, or in the worst of cases a 'political' matter, where there will be need for compromises - for giving and taking.

I hope that the utmost care will be manifested both in defining the objectives of the Chair and in finding a qualified person for the professorship. It goes without saying that the person considered for the chair should have very good scientific qualifications and, in addition, extensive experience of what the problems are in parapsychological research. Ideally, the person should have a background both within the natural sciences and in psychology. Beside being academically competent, the person should have the ability to cooperate and to make efficient use of available research facilities within the Department of Psychology. Furthermore, the person should be able to establish cooperation with other departments of relevance to parapsychological research. But besides being level-headed, scientifically competent, and flexible, the person should have an independent mind. The professor must be strong enough to stand pressure from all those parties who want to make the person an instrument for their own purposes.

It has been said, and with some justification, that the only clear-cut effect that we have in parapsychology is the so-called experimenter effect. Independently of how one may interpret the cause of such an effect, I find it important to pose the question: Should the person one considers for the chair be a 'psi-conducive' experimenter or not? Considering the uncertainties and vagueness of the concept 'psi-conducive', I think that one should follow a pragmatic line. That would be to try to find a parapsychologist who one could consider, based on empirical evidence, as standing somewhere between the two extremes. Should the professor be a 'sheep' or 'goat'? It would be farsical if some of the adamant critics were seriously considered for the Koestler Chair! As a contrast, I would not consider it as a real mistake if a critic of Ray Hyman's statue and inclination were considered for the task.

It is also to be hoped that the committee will pay due attention to the importance of getting a person of great integrity, for instance someone who is not prone to speak ill of or gossip about colleagues and critics, a malpractice or malady too often found in an academic setting. Albion may not be an exception in that respect.

It is certainly expected by responsible parapsychologists, and probably also by their critics, that Dr. Beloff will be given a say in the appointment of the chairholder.

Martin Johnson

EUR.J.PARAPSYCHOL., V.5, 1984

POTENTIAL PARANORMAL VALUE OF STATEMENTS OF PSYCHICS ACQUIRED UNDER FEEDBACK CONDITIONS

Hendrik G. Boerenkamp
University of Utrecht

Current methods of evaluating verbal statements made by psychics require that the target person score all statements for self-applicability. Some statistical evaluation procedures, for instance the Pratt-Birge method, even require that subjects score in addition the statements which were intended for other target persons. These methods suppose a data base made up either of correct statements only, or of all statements split up into correct and incorrect ones. A practical objection to these procedures is that the target persons are required to check large amounts of statements of which most are of a general nature.

Traditionally a paranormal character is attributed to a statement made by a psychic about a target person when two criteria are met: (1) a sufficient degree of correspondence between statement and reality and (2) no logical explanation of the correspondence. Current evaluation methods are based on all statements regardless of whether or not they fulfill the first criterion ('degree of correspondence'). Then the subsequent statistical procedure evaluates whether the statements meet the second criterion ('no logical explanation'). These methods are best suited for a model that assumes that all statements by a psychic might be influenced by psi.

In a previous paper (Boerenkamp and Schouten, 1983) a different approach was proposed, called the window model. The window model is based on the assumption that psychics are only occasionally able to use their psi ability and therefore statements of a general nature are considered as noise. Such statements are first eliminated based on ratings by judges. With this approach the data base is made up of statements selected for meeting the criterion of 'no logical explanation'. Only the retained items, the statements with potential paranormal value, have to be scored by the target persons for self-applicability, and these are the basis for a statistical test for paranormality. So actually the window model first reduces the number of statements by rejecting all statements of a general nature. The remaining statements with sufficient potential paranormal value are then sujected to the familiar procedures for evaluating verbal material.

A major advantage of the method is that it strongly reduces the number of statements which have to be scored by the target persons. Another advantage is that invalid estimates of probability of correspondence cannot yield spurious results as regards the significance of the experiment. The only effect of invalid estimates can be the unwanted inclusion or exclusion of statements.

In the previous study it was demonstrated that judges are able to estimate the potential paranormal value of statements to a sufficient degree of reliability and accuracy. However, the statements used in that study were acquired from a psychic in a 'no-feedback' condition. The psychic was presented a photograph of a target person who was unknown both to the psychic and the sitter. Consequently, the sitter was not able to provide any feedback about the psychic's statements. Although this is a common procedure in experimental research with psychics, it produces a constraint because in daily practice a sitter consults a psychic about a person the sitter knows and is interested in. Thus the normal situation is that psychics get feedback from the sitter about the statements which are made in the session.

Another limitation of the previous study concerns the fact that potential paranormal value was only judged by a probability of correspondence criterion. Pratt argued that the fact that much of the material from psychics is nonspecific does not say that it has no value from the standpoint of parapsychology: "Vague utterances may reflect more truly the manner in which psi finds overt expression in verbal material" (Pratt, 1969, p. 73). However, if some vague statements are potentially paranormal then we have to find another criterion, different from the probability of correspondence criterion, in order to distinguish between vague statements with a high potential paranormal value and statements with a low potential paranormal value. A suitable candidate for such a criterion seems to be the degree of spontaneity of the statement.

Lack of spontaneity is operationalized as the extent to which a statement probably is deduced from the information the psychic had already acquired or was given. Spontaneity is related to degree of deductive reasoning. A high degree of deductive reasoning implies that the content of the statement probably is derived from the foregoing verbal interaction. As in the case of a high probability of correspondence, a high degree of deductive reasoning is associated with a low degree of potential paranormal value. To a certain extent the criterion of probability of correspondence and the criterion of spontaneity are related to each other. Statements cannot have a low probability of correspondence if the content follows logically from available information. However, statements with a high probability of correspondence can reflect a low degree of deductive reasoning. For instance, if it is stated that the target person spent his previous vacation in a foreign country the probability for that statement in the case of a Dutch citizen is relatively high, but the degree of deductive reasoning can be nill if so far in the session nothing has been said from which anything could be deduced concerning the target person's vacation habits, feelings about foreign countries, opinions about travelling etc..

The aim of the present study was to investigate whether the applicability of the window model can be extended to statements obtained in a session in which feedback is provided, applying a criterion of probability of correspondence as well as a criterion of spontaneity. Hence it was investigated whether judges are able to give reliable estimates of the potential paranormal value of statements based on each of these criteria when the statements are obtained in a feedback condition. For the final selection of statements, a criterion was used which is based on both the estimate of probability of correspondence and the estimate of deductive reasoning.

Subjects

A distinction was made between psychics and mediums. The term psychic (or paragnost) is used for persons who believe themselves capable of intentionally giving information of paranormal value about a target person. A psychic is called a medium when he or she explicitly declares that such information is received from deceased persons. Of the 14 psychics who provided the statements on which this study is based seven can be considered mediums. The psychics are among the best known according to two Dutch parapsychologists, Tenhaeff and Zorab who published about most of them in the past, e.g. Tenhaeff (1960, 1962, 1965). The group of psychics consisted of 6 males and 8 females. They ranged in age from 30 to 75 with a mean of about 55 years. The mediums among them were all prominent members of the Dutch Society of Spiritists.

Eight judges participated in this study. These judges were volunteers, students taking part in a class on the evaluation of verbal material. They happened to be all males, approximately 25 years old.

THE STATEMENTS

The statements on which this study is based were obtained in sessions in which the sitter (Sybo A. Schouten) consulted 14 psychics about the fate of one of his friends who was missing at that time. The sitter had been requested by the wife of the target person to do so because she was greatly worried about her husband and feared for his life.

In the sessions the psychic was presented a photograph of the target person (man, about 35 years old) and was told that the target person was missing. The psychic was invited to provide as much information as possible about the present circumstances of the target person as well as to give other impressions concerning the person. The sitter answered requests for information (for example how long the person had been missing, if he was married, etc.) and corrected wrong statements. The sitter avoided giving additional information as much as possible. Under such conditions it might be assumed that the psychics behaved like they normally do when being consulted by clients.

ITEMIZATION

The sessions involving psychics were attended by the author as an observer. He recorded all sessions and typed them out in a standardized way. The 14 transcripts were ordered in temporal sequence in a 'book' consisting of 160 pages.

It was explained to the judges that the verbal utterances made in each session were to be split up into different statements according to specific rules. Judges were given training with a trial-feedback-consultation. Then, depending on the length of the session, each of the judges itemized one, two, or three sessions. Before the start of the class, it was decided randomly which sessions each judge would itemize. The itemization of each session was checked by another judge. Disagreements between the two judges about the way a particular part of a session should be itemized were discussed with the author and those parts were itemized by him.

Itemization was done according to the following rules:

(a) Statements made by a psychic about the target person were considered different when they conveyed information about different topics.

(b) A statement was defined as consisting of a psychic's verbal behavior about one of the topics on the topic-list. A statement was considered complete if any one of the criteria stated in rule 'd' applied. Verbal behavior of the sitter about one of those topics was called an 'informative action'. Statements made by the psychic and informative actions of the sitter were categorized in the same way by using the topic-list described in the following section.

TOPIC-LIST

Description of physical characteristics
110 Sex
111 Age
112 Appearance and overt behaviour (hair colour, way of walking)
113 Bodily health
114 Being alive or dead
Description of psychological characteristics

121 General personality traits (nervous, opiniated, submissive)
122 Temporal psychological circumstances (difficult period)
123 Religious and social orientation (catholic, conservative)
Description of relations
131 Relations with family members
132 Relations with friends, acquaintances, colleagues
133 Relation with sitter
Description of specific topics
141 Civil status, number of children, brothers, sisters
142 Profession, type of work, circumstances in work
143 House, type of house, circumstances in living
144 Leisure activity
145 Specific name (John) or property (car, cat) or event (accident)
Description of the situation related to the disappearance
151 Being alive or dead
152 Causes of the disappearance
153 Events from the moment of disappearance until the session
154 Developments in the case after the session

(c) The statements were listed and numbered sequentially in the order they appeared in the session. The informative actions were listed exactly in the same way. The topic of a statement was indicated on the list by putting the number of the category behind the statement in parentheses.
For example: The person is about 35 years old (111)
The first of the three digits indicates for whom the statement was intended. Statements about the target person are indicated by number 1. Psychics also make statements concerning persons related to the target person (e.g. wife, father, child). In that case the first number becomes 2.
For example: The mother of this person is about 75 years old (211).

(d) A statement was considered complete if a new statement was made in which a different topic was discussed, if the new statement concerned a different person, or if the statement was followed by an informative action by the sitter. Examples are:

(1) Another topic is discussed
Example: The person is very intelligent (121)
but ehh...he is not very healthy (113).
Excluded from this rule are verbalizations which are inclusive ('and'), exclusive ('or'), conditional ('if...then') or causal

('because'). In these cases both topics of the statement were represented in parentheses on the list.
Example: The person lives and works in Utrecht (143)(142)
A change in the temporal character (past, present, future) or the addition of advice or a warning was not considered as a change of topic. Such verbalizations were considered as inclusive or exclusive.
Example: He has a serious illness which he will get again (113)

(2) Another person is discussed
Example: The person is rigid (121)
ehh...his mother is a very nervous person (321).
This rule superseded rule (1) except for the excluded cases.
Example: He is rigid because his mother is anxious (121)(321).

(3) Informative action is provided by the sitter
Example: Psychic: The person is rather nervous and (121)
Sitter : Yes he is more or less (+?) (121)
Psychic: and a jealous type (121)
This rule superseded rule (1) and the excluded cases of rule (1).

(e) Informative actions of the sitter were broken down into confined and extended reactions. Confined reactions were defined as affirmation ('yes'); partial affirmation or affirmation with hesitation ('yes but'); unacquaintance (I don't know); partial denial or denial with hesitation ('no but'); and denial ('no'). These were indicated in the transcripts by the signs ++, +?, ??, -?, -- respectively. Included among the extended reactions were corrections of wrong statements made by the psychic and additional information given by the sitter.

(f) Repetitions and nearly repetitions were considered as separate statements if another statement or an informative action occurred between the original statement and its repetition, in accordance with rule (d).
Example: Psychic: The person works alone at his work (142)
Sitter : He does (++) (142)
Psychic: He almost always works alone (142)

NUMBER OF STATEMENTS AND INFORMATIVE ACTIONS

The number of statements made by the psychics was 902. The number of statements from each psychic ranged from 21 to 120 with a mean of about 64 statements.

The number of informative actions of the sitter was 558. The number of informative actions of the sitter varied from 15 to 102 with a mean of about 40 informative actions.

The seven mediums provided 478 statements, whereas the seven other psychics provided 424 statements. According to the Mann-Whitney-U test, the number of statements made by mediums proved not to be significantly different from the number of statements made by the other psychics. Five psychics made less than 40 statements, four between 40 and 80 and five made more than 80 statements.

METHOD OF ESTIMATING THE PROBABILITY OF CORRESPONDENCE

In the previous investigation (Boerenkamp and Schouten, 1983) judges were asked to estimate the probability of correspondence between the statements and the facts concerning the target person. They were requested to assign scores to each statement using the following criteria:

score 1: probability of correspondence is 50% or higher
score 2: probability of correspondence is about 33%
score 3: probability of correspondence is about 10%
score 4: probability of correspondence is 5% or less

In the present study the first task of the eight judges was of the same nature. In the previous investigation the judges assigned the scores to statements acquired in a no-feedback condition. However, in the present study the judges scored each statement on the basis of the information available to the psychic, including all informative actions of the sitter until that moment in the session. Presentation of the photograph of the target person to the psychic was rated as the first informative action by the sitter.

Example:

Sitter : (S1) We are here because a man is missing
(S2) This is a photograph of the person
(S3) He is a former friend and I have not seen him for ten years
Psychic: (P1) He is an easily distressed person

	(P2)	He has been missing several days already (?)
Sitter :	(S4)	It is seven days now...eight days perhaps
Psychic:	(P3)	Does he have relations in a foreign country (?)
Sitter :	(S5)	He has been abroad sometimes in connection with his work
Psychic:	(P4)	I would look in that direction...that he is wandering in a foreign country..either Germany or Belgium or England
	(P5)	He is easily distressed...very easily...

The task of the judge when estimating the probability of correspondence of P4 is to estimate the probability that the content of that statement is true 'considering the age, sex and appearance of the person, the fact that he is a former friend of the sitter, who has not seen him for the last ten years, the fact that he has been missing seven or eight days, and the fact that he sometimes was abroad in connection with his work'. In this example the estimation of the probability of correspondence was hardly effected by the feedback information. The information that the person had sometimes been abroad in connection with his work makes the option of 'wandering in a foreign country' a little bit less improbable. However, the probability of correspondence is still rather low. Therefore this statement might be assigned the score 4.

Before the judges actually scored the statements they took part in training sessions. The first part of the training consisted of scoring the statements of a trial-feedback-session which contained 34 statements and 21 informative actions. After independently scoring the first statement of this session the eight judges discussed the differences in their scoring. Then the second statement was scored and discussed and so on. In this way each judge could learn why in some cases his score differed from the group norm and this allowed him to adjust his manner of rating such items.

The second part of the training consisted of a group discussion about estimates of probability of possible facts in missing cases. For example, it was discussed which probabilities are associated with various causes for persons disappearing, how probable it is that a missing adult is still alive after being missing one day, after being missing a month, and so on. After the training, the eight judges estimated all the statements of the psychics independently.

RELIABILITY OF ESTIMATES OF PROBABILITY OF CORRESPONDENCE

In order to establish the reliability of the estimates the judges for each session were randomly split up into two groups of four judges each: i.e., judges 2,3,5,8 versus judges 1,4,6,7 for the session with psychic 1, judges 3,4,7,8 versus judges 1,2,5,6 for the session with psychic 2, and so on. For each group of four judges the average score for each statement of a session was established.

All 14 Spearman rank correlations between the average scores of the statements for the two groups were significant (rs varied from rs=.66, t=5.31, df=36, p<.01 to rs=.93, t=11.25, df=19, p<.01). Hence it is safe to conclude that a group of four judges can achieve reliable estimates of the probability of correspondence of statements from different psychics made in feedback conditions. After the scores assigned by the eight judges were summed all statements were classified in five categories of probability of correspondence. The observed distribution is presented in Table 1. We used a scale of five categories so as to make it comparable with scales used for the scores of 4 or 2 judges. The scores 14,18,22,26 were equally distributed over both adjacent categories.
Applying an arbitrarily chosen cut-off criterion between the medium and medium-high categories, the 14 sessions yielded 61 statements (7%) with low probability of correspondence. According to the Kolmogorov-Smirnov two-sample test, there was no significant difference between the distributions of scores presented in Table 1 of mediums and other psychics (Dmax=.079, Dcrit.05, two-tailed=.091).

METHOD OF ESTIMATING THE DEGREE OF DEDUCTIVE REASONING

In the task discussed above the judges concentrated on the content of the statement. In the present task the judges were asked to consider how many elements of information provided in the foregoing verbal interaction were included in the statement and to assign a score to each statement accordingly, using the following criteria:

TABLE 1
Distribution of scores of probability of correspondence

Probability of correspondence	high	hi-me	medium	me-lo	low
score	(8-14)	(14-18)	(18-22)	(22-26)	(26-32)
mediums	280	122	55	12	9
other psychics	215	115	54	22	18
all psychics	495	237	109	34	27
all psychics	55%	26%	12%	4%	3%

score 1: statement completely explainable by deductive reasoning; all informative elements in the statement were discussed before the statement was made
score 2: statement predominantly explainable by deductive reasoning; most informative elements in the statement were discussed before the statement was made
score 3: statement not predominantly explainable by deductive reasoning; most informative elements in the statement were not discussed before the statement was made
score 4: statement not explainable by deductive reasoning; none of the informative elements in the statement were discussed before the statement was made

Two general rules were used in assigning the scores. The first one was that repetitions or near repetitions of a statement were assigned a score of 1. The second rule was that whenever a statement was estimated as partly affected by deductive reasoning, the judge would count the proportion of deductive elements of information in the statement. If the number of deductive elements was greater than the number of new elements, the statement received a score of 2. For instance, in the example given in the paragraph 'Method of estimating the probability of correspondence', the task of the judges in scoring P4 would consist of estimating the degree of deductive reasoning in the statement. The judge would notice that the element 'foreign country' from the foregoing verbal interaction has been used but that the elements 'wandering around' and 'Belgium or Germany or England' are new. The statement thus might be assigned the score of 3.

For this task judges received training similar to that which preceded the estimations of probability of correspondence.

RELIABILITY OF ESTIMATES OF DEGREE OF DEDUCTIVE REASONING

The reliability of the estimates was established in the same way as in the case of the estimates of probability of correspondence. All 14 Spearman rank correlations between the average scores of the statements for the two groups of four (in random combinations out of eight for each session) yielded a significant correlation (rs varied from rs=.66, t=7.43, df=72, $p<.01$ to rs=.95, t=14.29, df=20, $p<.01$). By summing the scores assigned by the eight judges, all statements were classified into five categories in the same way as in the case of probability of correspondence. The observed distribution is presented in Table 2.

Applying an arbitrarily chosen cut-off criterion between the medium and medium-high categories, the 14 sessions yielded 95 statements (11%) reflecting a low degree of deductive reasoning (in other words, a high degree of spontaneity).

According to the Kolmogorov-Smirnov two-sample test, there was a significant difference between the distributions of mediums and of other psychics (Dmax=.138, Dcrit.01, two-tailed=.108). The mediums produced relatively more statements of a less spontaneous character.

THE POTENTIAL PARANORMAL VALUE OF THE STATEMENTS

As observed before, the two methods applied for estimating the

TABLE 2
Distribution of scores of degree of deductive reasoning

Degree of deductive reasoning	high	hi-me	medium	me-lo	low
score	(8-14)	(14-18)	(18-22)	(22-26)	(26-32)
mediums	242	122	76	28	10
other psychics	156	130	81	36	21
all psychics	398	252	157	64	31
all psychics	44%	28%	17%	7%	4%

potential paranormal value of each statement are to a certain extent related. The observed correlations between the average scores of probability of correspondence and of degree of deductive reasoning for two groups of four judges in random combinations appear to vary from rs=.21, t=1.99, df=82, p<.05, to rs=.64, t=8.94, df=118, p<.01). The correlations are primarily the result of the large number of nonspecific statements. Nonspecific statements were in general also judged as being low on spontaneity.

As observed before, applying a cut-off criterion between the medium and medium-high categories yielded 61 statements of low probability of correspondence and 95 statements of low degree of deductive reasoning. Of the 61 statements of low probability of correspondence, 37 (61%) also reflected a low degree of deductive reasoning. However, an insignificant correlation was observed between the average scores of probability of correspondence and the average scores of degree of deductive reasoning of the 119 statements which met at least one of both criteria (rs=.11; t=1.21, p=.23, df=117).

Applying both criteria, the potential paranormal value of statements may be expressed by the combined scores of probability of correspondence and degree of deductive reasoning. These combined scores appear as reliable if not more reliable than the individual scores. After splitting up the judges into two groups of four as described above, it appears that all 14 correlations for the groups based on these total scores are highly significant (rs varies from $rs=.73$, $t=8.94$, $df=72$, $p<.01$ to $rs=.94$, $t=12.25$, $df=20$, $p<.01$).

Applying the combined score as the criterion of potential paranormal value has the effect that most repetitions are excluded from the data base but that statements of medium probability of correspondence are included if the statement has a strong spontaneous character. Using 8 judges, the combined scores of statements range from 16 to 64. Applying a cut-off criterion between the medium and the medium-high categories, the criterion score is 44. A repetition is excluded from the list if a high score of probability of correspondence (for example 29) is added to a low score of probability of deduction (for example 11). A statement of medium probability of correspondence is included if a score of for example 20 is added to a high score of probability of deduction, for example 29. Classifying the combined scores in five categories as before, the distribution of potential paranormal value of statements is presented in Table 3.

It may be concluded that, applying a cut-off criterion between the medium and medium-high categories, 65 statements (7%) are rated as of sufficiently paranormal value to be included in the data-base for further analysis. This actually implies a reduction in number of statements from 902 to 65 (93%). The distributions observed for mediums and other psychics differ significantly ($Dmax=.133$, $Dcrit.01=.108$ two-tailed). Mediums made relatively more often statements of a general nature. However, both groups produced a number of statements of potential paranormal value.

NUMBER OF JUDGES REQUIRED FOR ESTIMATING POTENTIAL PARANORMAL VALUE

Even with the window model approach, the evaluation of a large number of statements from different experimental conditions would take much time if 4 judges were needed to evaluate each statement. Therefore it was decided to investigate whether 2 judges would yield sufficiently reliable estimates. For each session two groups of two

TABLE 3
Distribution of scores of potential paranormal value based on the added scores for probability of correspondence and degree of deductive reasoning

score	low (16-28)	lo-me (28-36)	medium (36-44)	me-hi (44-52)	high (52-64)
mediums	257	151	47	16	7
other psychics	176	143	63	26	16
all psychics	433	294	110	42	23
all psychics	48%	33%	12%	5%	2%

judges out of the eight judges available were randomly selected. For both groups, the average scores representing the potential paranormal value of the statements based on the combined scores for probability of correspondence and degree of deductive reasoning were compared. It turned out that all 14 correlations between the scores for both groups of two judges were almost as high as those between the average scores for two groups of four judges (rs varies from rs=.63, t=6.82, df=72, p<.01 to rs=.94, t=10.68, df=19, p<.01). When for groups of two judges the average scores of probability of correspondence and the average scores of probability of deduction were compared separately, the correlations were also highly significant. In the case of the estimates of probability of correspondence, rs varied from rs=.49, t=3.40, df=36, p<.01 to rs=.88, t=8.21, df=19, p<.01. In the case of estimates of probability of deduction rs varied from rs=.45, t=4.24, df=72, p<.01 to rs=.83, t=10.46, df=49, p<.01.

A comparison between the distribution of the 902 statements based on the estimates of potential paranormal value of eight judges and the distribution based on the estimates of two judges (randomly chosen from the eight judges for each session) yielded a nonsignificant difference (Dmax=.037, Dcrit.05, two-tailed=.064). The distributions are presented in Table 4. The same holds when the distribution of estimates of probability of correspondence and the distributions of estimates of degree of deductive reasoning were compared separately.

TABLE 4
Distributions of estimates of potential paranormal value based on the added scores of probability of correspondence and of degree of deductive reasoning for eight and for two judges

category	low	lo-me	medium	me-hi	high
Score based on eight judges	(16-28)	(28-36)	(36-44)	(44-52)	(52-64)
distribution	433	294	110	42	23
Score based on two judges	(4-7)	(7-9)	(9-11)	(11-13)	(13-16)
distribution	467	265	109	34	27

The relationship between the scores of two judges was also studied. From the eight judges for each session one pair of two judges was randomly selected. Then for each statement there were two scores of probability of correspondence, ranging from 4 to 1, and two scores of degree of deductive reasoning. These scores can be tabulated in two 4x4 matrices as depicted in Figure 1.

SCORE JUDGE 1

	4	3	2	1	
	A	B			4
SCORE	C	D			3
JUDGE 2			E	F	2
			G	H	1

FIGURE 1
The relation between the scores of one pair of judges

If the two judges agreed perfectly about each statement, all scores would be tabulated in the cells on the diagonal 4,4 to 1,1 (cells A,D,E,H). It turns out that 81% of the scores of probability of correspondence are found in the cells A,B,C,D,E,F,G,H. This means that two judges both estimated 81% of the statements as being either rather specific (score 4 or 3) or rather vague (score 2 and 1). Within the A,B,C,D cells, 69% of the scores are found in the A and D cells. Within the E,F,G,H cells, 58% of the scores are found in the E and H cells. From this it may be concluded that the judge's ability to estimate consists primarily of discriminating between the specific and the vague statements. A similar pattern is found in the case of degree of deductive reasoning. It appears relatively easy to discriminate between a score of 4 or 3 versus a score of 2 or 1 but difficult to discriminate within these groups. The type of relation between the scores which represents the potential paranormal value is established by combining the figures of both matrices. These figures are represented in Table 5.

Hence in the case of two judges, approximately 80% of the statements will receive the same classification as regards reflecting or not reflecting potential paranormal value.

TABLE 5
Relation between scores of potential paranormal value

	prob. of corresp.	prob. of deduction	potential par. value	potential par. value %
N ALL	902	902	1804	
N ABCDEFGH	734	738	1472	82%
N ABCD	52	93	145	
N AD	36	65	101	70%
N EFGH	682	645	1327	
N EH	393	429	822	62%

THE REAL PARANORMAL VALUE OF THE STATEMENTS

In this study the application of the window model resulted in reducing the number of statements from 902 to 65. These 65 statements are listed in the Appendix of the article. It is beyond the scope of this article to discuss the data extensively. However, some conclusions as regards the real paranormal value of these statements can already be drawn in this phase of the study.

The first one is that the internal consistency between different psychics appears rather low. Thus the statements of one psychic are often in conflict with the statements of another psychic. All possible causes for the target person being missing are offered by the various psychics. The same applies to the place where the target person was supposed to be found.

From the above it also follows that a great number of these conflicting statements with potential paranormal value must be wrong. Hence few statements of potential paranormal value are left which might be correct. Thus it can be concluded that in this case the 14 psychics yielded only a few statements with a reasonably low probability of correspondence and degree of deductive reasoning. None of the statements led to a recovery of the target person.

ABSTRACT

Current procedures of evaluating verbal statements made by psychics require that the target persons rate all statements for self-applicability and that all statements be included in the analysis. In general, the majority of the statements are rather vague in nature. In a previous study, a procedure was proposed in which statements are first classified by judges, not the target persons, into categories of degree of probability of correspondence. Then for the final evaluation, the data base is made up only of statements of sufficiently low probability of correspondence. It was demonstrated in that study that such a window model can be applied and that judges are able to give reliable and accurate estimates of probability of correspondence for statements acquired under non-feedback conditions.

In the present study, it was investigated whether the applicability of the window model could be extended to statements made by psychics in sessions in which a sitter gives feedback. Fourteen psychics provided 902 statements when consulted about the fate of a person who had disappeared. It was studied whether judges are able to give reliable estimates for probability of correspondence for these statements. In addition, judges were also requested to give estimates for degree of deductive reasoning. A high degree of deductive reasoning implies that the informational content of the statement could have been deduced from the foregoing verbal interactions in the session between psychic and sitter.

The data indicate that pairs of judges are able to give reliable estimates of probability of correspondence and of degree of deductive reasoning for the statements. Application of a criterion based on a combination of the two estimates resulted in a 93% reduction in the number of statements for the final evaluation (from 902 to 65).

REFERENCES

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

Burdick, D.S., Kelley, E.F., 'Statistical methods in parapsychological research' in Wolman, B.B. (Ed) 'Handbook of Parapsychology', van Nostrand Reinhold, 1977, pp 81-131.

Millar, B., 'The observational theories: A primer', E.J.P.,1978,2, 304-322.

Pratt, J.G., 'On the evaluation of verbal material in parapsychology', Parapsychological Monographs No 10, New York, Parapsychology Foundation, 1969.

Tenhaeff, W.H.C., 'Proceedings of the Parapsychological Institute of the State University of Utrecht, number 1 (1960), number 2 (1962), number 3 (1965).

APPENDIX

Statements with potential paranormal value

Medium 1:

1. He works in a high building. In the neighbourhood of the building, flats are under construction. It is in the surroundings of Amsterdam.
2. A night watchman is involved in the case, either as a bystander witnessing an accident in which the missing person was involved or as a bystander witnessing the missing person being held up, or as the one who held up the missing person.
3. In his work the person wears a white coat.
4. When the person who held up the missing person is caught, he will be in the possession of the pocket book, watch, and driver's license of the person.
5. A car involved in the case, either belonging to the person who held up the missing person or to the missing person, has the registration number 63-14-DH or 63-14-HD or DH-63-14 or HD-63-14 or DH-17-34 or HD-17-34.
6. The person who held up the missing person is a swindler of cars.

Medium 2:

7. The person has drowned.
8. The person has or had a moustache.
9. The person has drowned in the port of Rotterdam, and is hooked onto a boat.

Medium 4:

10. The person is not really missing.
11. He is left handed or has trouble with his left hand.
12. In his childhood he was dominated by a tall heavy man.
13. A person or dog in his surroundings is called Bas.
14. He has a very particular type of religious devotion.
15. The person has a dent in his skull at this moment. It is a consequence of either an accident in his youth or of having fallen into something or of being smashed in, in the time from the moment of being missing until now.
16. The person has something to do with a slipway.
17. The name of the person is Wim.

Medium 5:

18. The person is troubled by struma.
19. The person lives close to a wood.
20. The person was married seven years ago.
21. Something is wrong with the backside of his head as a consequence of an accident.

Medium 6:
22. The person feels strongly attracted to water when he is in a depressed mood.
Medium 7:
23. The person was not on good terms with his wife and he was in love with another woman.

Psychic 8:
24. The person is wandering somewhere in a foreign country, either Germany or Belgium or England.
25. The sitter knew the person as a psychiatric patient in the years when they saw each other frequently.
Psychic 9:
26. The person was born in the province Limburg.
27. The person will give a sign of life on the tenth day from today, on Saturday.
28. A large bridge is visible at the left of his house.
29. There is a sports field in the neighborhood of his house.
Psychic 10:
30. The person's wife plays tennis.
31. The person has several cats in his house.
32. His father acts very carefully and carries out everything according to plan and makes annotations in books.
33. The person's father lives outside the built-up area of a city or village.
34. The intersection 'Quatre-Bras' in the neighborhood of the village Austerlitz has something to do with the missing case or the person lives in that area.
35. He lives in the area of the village Bloemendaal.
36. The person has gotten a promotion very recently.
37. The person's car has a red colour.
Psychic 11:
38. (See drawing). The person drowned in a kind of canal flowing into the sea. Drawing elements A: sea, canal, boat, bridge, distance boat-bridge, church or other high building, factory or large house, broken window in the facory or large house.
39. The person has been missing fourteen days.
40. (See drawing). Addition of B elements: truck, hotel, large white plate with black characters in front of the hotel.
41. (See drawing). Addition of C elements: place of drowning and place of the dead body.
42. The person had unknown liabilities and that is the most important cause of the suicide.

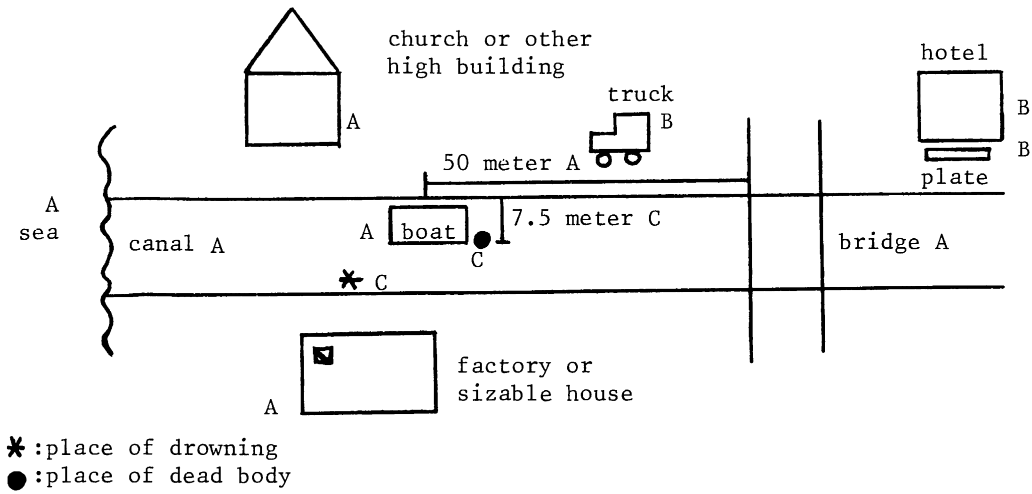

FIGURE 2
Drawing from psychic 11

43. The person has two little daughters .
44. The person was bothered by his heel.
45. The person was bothered by his right heel.
Psychic 12:
46. The person did not get enough love in his youth and was raised very conservatively; rigid parents or rigid grandmothers frustrated him in such a way that he has never succeeded in throwing off that yoke.
47. The person is or was alone somewhere in a neither very high nor very low house, on the first floor.
48. The person had homosexual tendencies, for example he was being sexually attracted by his master. He could not satisfy these needs as he was married.
49. Another person in the family, in the family of his father has

committed suicide.
Psychic 13:
50. The person was shocked once by the fact that a family member suddenly committed suicide.
51. The person had already attempted suicide in the past.
52. The person is good-natured but mentally ill because his mother was mentally ill too.
Psychic 14:
53. The person is 27 years old (before seeing photograph).
54. The person has been missing since Wednesday a week ago.
55. The person is over-excited and alive.
56. He is missing because he is in love with another woman.
57. The sitter did not like the missing person at all in the years they saw each other frequently.
58. If you were to visit the person at home you would have to go by train from Utrecht to Nijmegen, and you would have to transfer there to another train.
59. The person's wife has an idea of where he is, as the person took his passport and clothes with him.
60. A few days before becoming missing he visited a doctor.
61. The person's wife is 29 years old.
62. The person is alive and his car has not yet been found.
63. The person recently bought new glasses.
64. The person visited a psychic just before he became missing.
65. After parking his car, the person turned to the right into a park with water and he drowned in that water.

EUR.J.PARAPSYCHOL., V.5, 1984

IS THERE A PARANORMAL (PRECOGNITIVE) INFLUENCE IN CERTAIN TYPES OF PERCEPTUAL SEQUENCES ? PART II

Holger Klintman
Lund University

STUDY V
EXTENSION OF TIME RANGE

BACKGROUND AND HYPOTHESIS

In the four experiments described in Part I (Klintman, 1983a), the time distance between the first event at which TRI (Time-Reversed Interference) was observed and the source of the interference (the second event) ranged from .8 to 3 seconds. Also, in all these studies there seemed to be important individual differences with regard to the conditions critical for the occurrence of the interference: it was suggested that in some subjects, 'facilitators', identical or congruent events were a necessary condition, while in others, 'inhibitors', the two successive events would have to be dissimilar or incongruous for the interference to take place. And further, in Study I there was evidence that these conditions for precognitive interference were in fact similar to those pertaining to ordinary proactive cognitive interference.

This apparent correspondence between the two types of interference had an interesting practical implication: as there was evidence that individual differences in terms of facilitation/inhibition of a second cognitive event as a function of a first were confined to interstimulus intervals up to about 4 seconds only (Klintman, 1983b), the same might be true of TRI. Accordingly, if the time interval between the two events was increased from the 3 seconds of Studies II-IV to above the 4 seconds level, then the critical relation between the meanings of the two events would tend to be similar for all subjects. Specifically, if the interstimulus interval were set to 6 seconds, the necessary condition for TRI would be non-identical events (Klintman 1983b). This was one of two hypotheses tested in the present study. According to our basic assumption, this would result in an observed predominance of identical pairs of (odd/even) events.

In addition, since in all the preceding studies the TRI had been observed only in sequences where the first comparison yielded nonequality (odd/even or even/odd), the same condition was expected to obtain in the present experiment.

The second hypothesis concerned the time interval between the earliest observable TRI effect and the target event. Could the total time be increased from a few seconds to, say, 10 minutes? To answer this question the following hypothesis was tested:

> If a cognitive event E at time t(1) is repeated at time t(2) > t(1) and at that time subjected to TRI, then the interference will also affect the event at time t(1).

This hypothesis is essentially implied in the conception of a spreading effect, discussed in Part I.

The design of the experiment was based on that of Studies II-IV, using selective temporal feedback as a means for detecting the time shifts. However, in the first event both of the one-digit numbers were printed on cards, the cards being drawn from two randomized sets. As before, in the second event the microsecond digit between 0 and 9, produced by the subject in an estimated interval of 1-2 seconds on the digital counter display, was compared with the number printed on the card (drawn from a third randomized set). The time between the two events was 10 minutes.

TIME MODEL

Selective temporal feedback (STF)

STF may be defined as the change of the first event, E1, into E1′ at the same point in time, t(1), the change taking place if, and only if,

1. E1 has a specific meaning M, and
2. a second specified event, E2, will occur at a later point in time, t(2).

It may be helpful to visualize the STF and other TRI systems in a two-dimensional time plane, where the horizontal axis (t) is 'ordinary' time, and the y-axis is the second time-dimension (τ ,tau). At any point t(x) we may represent two or more events as points on a vertical line extending into τ , the events at the top being the 'actual' or 'observable' events and the lower ones being 'conditional' events (see figure 1).

In figure 1 open arrows (——▷) indicate forward causation while closed arrows (◀——) indicate 'backward causation', the course of TRI. As shown in the figure, the event E2 in t(2) is dependent on the occurrence of a specific event E1 in t(1). According to the definition above, in our example, the existence of E1 constitutes the necessary and sufficient condition for STF. The feedback is indicated by the closed arrow from right to left, and the associated change of E1 into E1′ by the closed arrow between E1 and E1′.

The spreading effect

As a further example of two-dimensional time representation, the causal relations involved in the production of a spreading effect (see also Part I) are shown in figure 2.

As the reader may remember, the conception of multidimensional time as a means for better understanding the apparent paradoxes of precognition is not new in the history of parapsychology. Perhaps the best known proponents of such systems are Dunne (1927) and Broad (1962), the latter sketching an interesting basis for further elaboration. The simple model used in the present context should, of

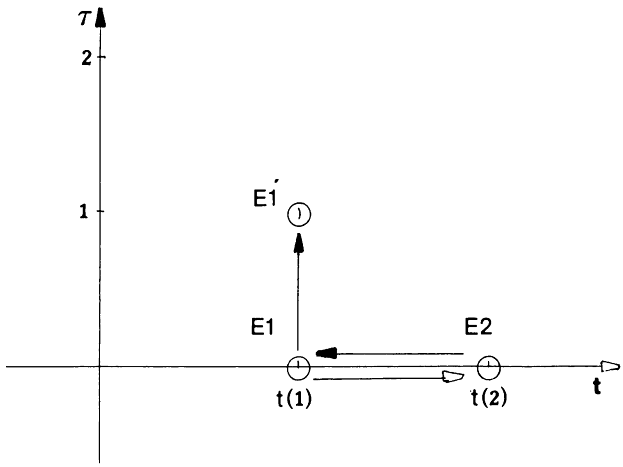

FIGURE 1
Representation of forward and backward causation in two-dimensional time.

course, not be seen as a complete system for temporal causation, but rather as the specification of one particular aspect of such a hypothetical system, selected and defined for the purpose of description and empirical testing.

METHOD

Procedure

The experiment consisted of two sessions (I and II) separated by a 10 minute break. The subjects' task was to make three comparisons or matches (C1, C1′ and C2) of two one-digit numbers with regard to the odd/even property. A 'trial' consisted of two parts, the first taking

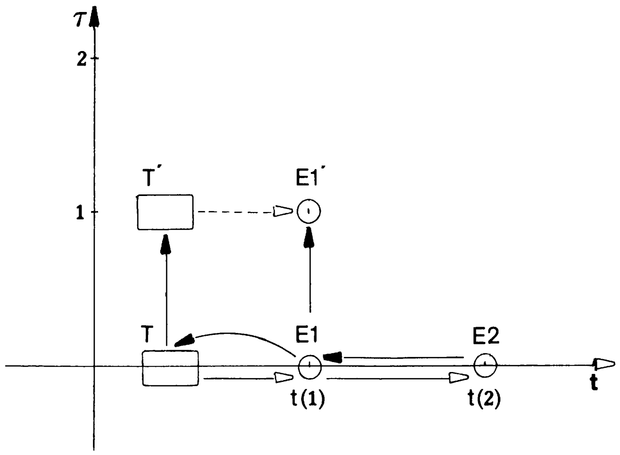

FIGURE 2
The Spreading Effect. TRI primarily acting from E2 to E1 spreads to decisions associated with the subject's production of time interval T, causing a change in the length of the latter (T′). The broken arrow represents the closing of a feedback loop.

place during session I, the second during session II, i.e., after a ten minute pause. Match C1 was made in session I, C1′ and C2 in session II. Match C1′ was always identical to C1 and included a rematching of the numbers on two cards from randomized sets, whereas C2 was a matching of a number on a card from a third randomized set with a number between 0 and 9 produced in session I on the digital counter (1 microsecond time-base). A series of six trials was given, the trials being so arranged that their respective first parts succeeded each other before, and their second parts after the ten minute pause.

Figure 3 shows the temporal patterning of the experiment. In the figure, the subject's identification of the numbers on the respective sets of cards is designated by S1, S2, and S3, a match of two such

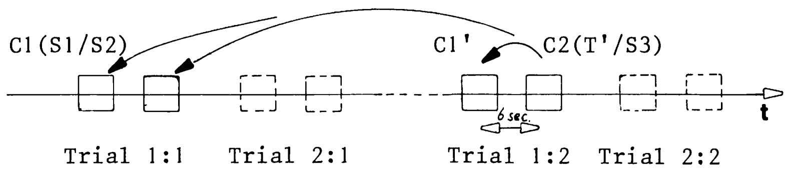

FIGURE 3

Temporal design of the experiment. One trial is shown, the place of a second indicated. The arrows from right to left indicate the path of TRI. Other symbols are explained in the text.

numbers by C(Si/Sj), and the produced time interval by T. T′ is the subject′s record of T as made in Part I, and C2(t′/S3) is the matching of this record with the set 3 card in session II. Stimulus material and apparatus were identical with those used in Studies II, III, and IV (see Part I). For further details of the procedure, see Appendix I.

Subjects and Experimenter

Sixty-five subjects (university students, unselected sample) participated in the experiment. As in earlier studies, double-blind conditions were used in order to minimize experimenter/subject bias. Again, in this study a different experimenter was used.

RESULTS

According to the null hypothesis, i.e., under purely chance conditions, the probability would be the same, 1/4, for each of the four possible combined outcomes of the first and the last matches (nonequal/equal, equal/nonequal, equal/equal, and nonequal/nonequal, cf. Part I). However, the results indicated a statistically significant deviation of the response patterns from such uniformity. Table 1 shows the TRP (for a definition, see Part I) frequency distributions over the four combinations of C1 and C2. Entered in the table are the number of subjects with the respective TRPs, the subdivision into (a) and (b) being the same way as in the former studies.

TABLE 1
Frequency distribution of typical response pattern (TRP) when C1 yielded nonequality (ne) and C2 nonequality or equality (e) and when C1 yielded equality and C2 equality or nonequality (compare table 9, Part I)

	TRP			TRP		
	ne /ne 1 2	ne /e 1 2	=	e /e 1 2	e /ne 1 2	=
	38	19	(8)	27	28	(10)
(a)	P=.008 (Binomial test)			(b) P>.20 (Binomial test)		

As predicted, in the undifferentiated group a predominance of TRPs of identical events was observed, this being the case only in sequences where the first match yielded nonequality (p=.008, one-tailed test).

Hence, the hypotheses regarding (1) uniform TRI conditions in sequences with a 6 seconds interstimulus interval, and (2) the feasibility of an extension of the total time for the observation of the effect, were both confirmed.

ABSTRACT

Study V reported another experiment of precognitive or time-reversed interference (TRI) in perceptual sequences. Study V represented a cross-validation but also an extension of the scope of previous studies in the series. The sample consisted of 65 unselected university students. The experiment tested two hypotheses based on findings from an earlier series of studies (Part I) by the same author, viz. concerning (1) the possible extension of TRI to cases where the sequential events were separated by intervals of up to ten minutes, and (2) reduction of inter-individual variation in such systems. The method employed a temporal feedback design, and the interacting events were numerical tasks and time estimations. The results confirmed both hypotheses at a level of statistical significance and were interpreted within the framework of a two-dimensional time model. Also, the joint results of the present and of the preceding studies were discussed and some general conclusions drawn.

CONCLUDING DISCUSSION: STUDIES I-V

The five studies presented here were concerned with the problem of precognitive interference in sequences of perceptual events. Prompted by casual observation of some seemingly anomalous relations in ordinary experimental work with repeated/non-repeated stimuli, the question originally asked was whether the meaning of a non-inferred second perceptual/cognitive event could affect the time-course of a first. Apparently, this hypothesis called into question the 'irreversability of causation', and before any serious attempt could be made at explaining the observation along these lines two conditions had to be fulfilled: (1) that the observations were not caused by purely chance variation, and (2) that no artifacts were involved.

This required a series of experiments in which previous results were replicated using a design which excluded or else controlled for potential sources of artifact. This was attempted with some success in the series summarized in the present articles.

Looking back at the results of the studies, some general characteristics of TRI in short-term sequences suggest themselves.

Some of the more important are listed in the following:

1. The effect is contingent upon the relation of meaning between the two events.

In Study I the change in RT to the first event (E1) was different when E1 was followed by a second event (E2) with the same meaning as compared to the case of an incongruent E2. In Studies II-V the meaning of the second match (C2) in relation to that of the first (C1), identical or non-identical, was the differentiating factor.

2. The importance of individual differences is evident, notably regarding the individual's tendency to inhibit or, alternatively, facilitate perceptual/cognitive activity in the afterphase of an initial percept. Thus in Study I, subjects' individual types of forward interference (positive or negative) were predictive for the sign of their backward (time-reversed) interference. Also, in Studies II-IV, still using a between-stimulus interval of less than 4 seconds, the subjects' negative afterimage duration times were predictive of the inter-event conditions critical for the TRI effect.

In the literature there is abundant evidence of the importance of personality variables for ESP performance. For example, in several studies a relationship between introversion/extraversion and ESP was indicated (Humphrey, 1951; Nicol & Humphrey, 1953; Kanthamani & Rao, 1972). This lends some support to the present findings regarding the relationship of TRI with visual aftereffects, since facilitation of internal perceptual activity, as evidenced after long afterimage durations, is frequently one characteristic of the introvert personality as contrasted to the extravert (Andersson et al., 1972).

3. TRI tends to be stronger in the top half than in the bottom half of a series of trials. This was repeatedly observed in Studies II-IV, which all included two successive 5-trial series (series a and b). In each experiment, the effect of TRI yielded much greater contrasts in series a than in series b.

In fact, these kinds of top/bottom as well as quarter distributions have been frequently observed in ESP experiments (e.g., Pratt, 1944; Rhine & Humphrey, 1944). It has been suggested that such differences may be associated with a gradual reduction in the 'novelty' of the task as the series progresses (e.g., Pratt & Woodruff, 1939; Hallet, 1952). Incidentally, I have regularly observed a similar decline in the strength of ordinary 'forward' interference of the meaning of a first stimulus with the identification of a second one, the most marked effects being present when the task is fresh.

Mention should also be made of the fact that in Studies II-V TRI was observed only if the first event yielded NON-equality. The reason for this might be that the non-equality relation, as compared to the equality relation, produces the more durable residual activity in the afterphase of the original cognition, and that a minimum level of such activity is required at the time of E2 for TRI to occur.

4. The phenomenon is present in 'ordinary' samples of university students.

5. The phenomenon is relatively insensitive to the choice of experimenter. Since in each study a different experimenter served, the presence of TRI was probably not dependent on some idiosyncratic experimenter characteristic. Indeed, the experiments were carried out under rather routine conditions, without any long-term buildup of the subject-experimenter relation.

6. The phenomenon has some amount of retest stability. In study IV, subjects from Studies II and III were retested after about six months, yielding similar results to those in the previous sessions.

7. TRI may cause RT changes of up to at least 65 milliseconds in the first event. (Study I).

8. The TRI of E2 with E1 tends to spread to events (E) prior to and associated with E1. The spreading effect is evidenced in a disturbance (change) of the cognitive activity included in E (in our case subjects' decisions to start and stop the timer).

9. Negative feedback loops, involving proactive interference ('ordinary' causation) and time-reversed interference ('backward' causation) may be feasible. Studies II-V were based on this assumption, and its consequences for the frequency distributions of response patterns were successfully predicted. A two-dimensional time system was proposed as a heuristic device for the description of such loops, postulating a step-wise increase in time (τ) as an event is changed by TRI in time (t).

A noteworthy advantage of detecting TRI by means of negative feedback techniques rather than by using central statistics methods is the fact that the former allows detection of far smaller time changes than does the latter. In fact, the maximum sensitivity of the feedback design approaches that of the resolution of the counter. Also, by setting the time base of the counter, the experimenter has a means of varying at will the sensitivity of the measurement.

A reference to the observational theory

Looking now at the current trends in parapsychological theory, I am struck by the similarity between some aspects of observational theory (see Millar, 1978) and the model of TRI and selective feedback adopted in the present article.

The concept of feedback is at the center of observational theory: for example, under certain conditions, if at time t(1) a subject is free to choose between two equally probable alternatives in order to guess which one will occur at time t(2), then, according to the theory, his choice may be affected by the feedback he would experience at time t(2). It is assumed that the subject's choice behavior may be ultimately contingent upon an inner (neural) RNG, the state of which may be altered by feedback of the outcome of his choice. It is further suggested that such feedback may act selectively to produce a more desired or pleasurable result.

In the present experiments, feedback seems indeed to be a necessary part of the system. However, in our interpretation, no purposeful selective mechanism is called for and no within-subject feedback in terms of right/wrong or pleasant/unpleasant is involved. Rather, feedback - or the matching of a first event (E1) with a second (E2) - follows from cognitive interaction between the meaning of E1 ('equal' or 'non-equal') and that of E2 ('equal' or 'non-equal'). If we assume that at t(1) the precise time (to the millionth of a second) for the subject's decision to stop the timer is ultimately a function of the uncertainties of his inner RNG, it seems plausible that under some conditions the RNG might be affected by the subsequent cognitive match resulting from this decision. The selective nature of such a system would follow if only one of two possible matches was associated with feedback interference. (In the present experiments, there were indications that the type of match conducive to feedback interference is one of equivalent meanings of E1 (or its residual) and E2).

The hypothetical model sketched in this article involves a system relatively insensitive to extra-subject interference. From a practical point of view, then, such a system may be regarded as closed, the factors mainly responsible for the TRI being similarities and contrasts of contiguous cognitive events. In fact, as pointed out earlier, the conditions for time-reversed interference would be strikingly similar to those of 'ordinary' proactive cognitive interference. Clearly, from the point of view of observational theory, the question may be raised as to whether the subject's observation of his own time-estimates in Studies II-V is at all relevant to the outcome: might not his inner RNG, which triggers the decision to stop the timer (or, alternatively, the timer itself) be directly affected by, say, the author's ultimate appraisal of the results? In principle, this would imply an open system, where the estimated time intervals, and so the experiment as a whole, might be at the mercy of a host of feedback agents, some perhaps more potent than others, at different points in time. For the present experiments, the consequence of this state of affairs would probably be difficulties with the replicability by different investigators. Conversely, a closed system would very likely insure a high degree of replicability.

At least one experiment suggests a differentiation between the two types of systems, viz. one in which, other conditions being equal, the digital output of the timer (unknowingly to the subject) is geared to an external random time generator. If under these conditions the external RNG yielded results similar to those of the present studies, an open system would be indicated. However, the test would have to be taken further in order to demonstrate that the level of replicability was the same for the two contingencies. A follow-up using instead a rigid (pre-programmed) number generator would then constitute a further test of the role of the external RNG. If on the other hand the statistical effects in the present experiment would vanish (or at least the level of replicability be reduced) by the connection of an external RNG, then the validity of an interpretation of the results in terms of a mainly intra-individual system would be strengthened. This would indicate that the cognitive interaction in the stimulus sequences used added to the subject's behavior some systematic variation which was not altogether overridden by extra-experimental psi sources.

In general, the problem of 'delimiting' a psi-system down to a relatively few and well-defined agents and relations seems to be a crucial one. A somewhat different approach would be to first set up a system where psi effects are created of a strength well above that of possible extra-system agents, and then to attempt to select intra-system effects for detection, perhaps by using a detection technique which incorporates variable sensitivity. While this may sound simple enough, one would be wise not to underestimate the possible practical - and theoretical - complications of this type of research.

REFERENCES

Andersson, A., Nilsson, A., Ruuth, E., Smith, G. 'Visual aftereffects and the individual as an adaptive system'. Lund: Gleerup, 1972.

Broad, C. D. 'The notion of "precognition"', In J.R. Smythies (ed.), Science and ESP, New York: Humanities Press, 1967.

Dunne, J.W. 'An experiment with time', New York: Macmillan, 1927.

Hallet, S.J.A. 'A study of the effect of conditioning on multiple-aspect ESP scoring', J.o.P., 16, 1952, 204-211.

Humphrey, B.M. 'Introversion-extraversion ratings in relation to scores in ESP tests', J.o.P., 15, 1951, 252-262.

Kanthamani, B.K., Rao, K.R. 'Personality characteristics of ESP subjects: V. Graphic expansiveness and ESP', J.o.P., 37, 1973, 119-129.

Klintman, H. 'Perceptual sequences and recognition level', Psycholog. Res. Bull., Lund University, 23, No 8, 1983a.

Klintman, H. 'Is there a paranormal influence in certain types of perceptual sequences: Part I', E.J.P., 5, 1983b, 19-49.

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

Nicol, J.F., Humphrey, B. 'The exploration of ESP and human personality', J. A.S.P.R., 47, 1953, 133-178.

Pratt, J.G., Woodruff, J.L. 'Size of stimulus symbols in extrasensory perception', J.o.P., 3, 1939, 121-158.

Pratt, J.G. 'A reinvestigation of the quarter distribution of the (PK) page', J.o.P., 8, 1944, 61-63.

Rhine, J.B., Humphrey, B.M. 'The PK effect: Special evidence from hit patterns. II. Quarter distributions of the set', J.o.P., 8, 1944, 254-271.

Holger Klintman
Department of Psychology
Lund University
Paradisgat 5
S-22350 Lund
Sweden

APPENDIX

Instruction

The following instruction was given each subject verbally before the experiment started.

"The device on the table in front of you is an electronic timer measuring the length of the time elapsing between two events. If you press this button" (indicating) "the timer starts, and when you press this one the timer stops counting. During the time between two such events the numbers here on the display run through the series of 0 to 9 over and over again. The series will stop at one of these numbers as you press the 'stop' button.

The procedure will be as follows: When I say 'DIGIT' you are to press the 'start' button. Then, after about 1 to 2 seconds you are to press the 'stop' button and check the resulting number by sliding this little screen to the left, like this" (demonstrating). "After checking the number, replace the screen and report the number to me. Also record the number here in your protocol. Now, when you check the numbers it is important that you decide whether they are odd or even (1,3,5,7,9 are odd; 2,4,6,8,0 are even numbers).

Also in front of you, in the stands, there are three sets of cards, each with a one-digit number printed on it" (demonstrating a test card. In the stands, the back of each card is facing the subject). "Card 1 of set 1 goes together with card 1 of sets 2 and 3, card 2 of set 1 with card 2 of sets 2 and 3, etc. When I say 'CARDS' your task will be to simultaneously turn over the first cards of sets 1 and 2 and decide whether their numbers are the same or different with regard to being odd or even. If, for instance, one number is 5 and the other 9, then these numbers, both being odd, are classified as 'same'. If one is 2 and the other 7, then they are 'different', since one is even and the other odd. As soon as you have compared the numbers, place the cards on the table, face down, and report to me 'same' or 'different', depending on the result of the comparison you've just made. Then, after about three seconds, I will say the word 'DIGIT'. This signals you to produce a number on the timer, read it out aloud, check whether it is odd or even, and record the number in your protocol. After that there will be a 15 sec. pause, followed by my saying 'CARDS' again, which marks the beginning of the next trial. Then turn over the next pair of cards of sets 1 and 2, compare them with respect to their being odd or even, report to me 'same' or 'different', and then, when I say 'DIGIT', produce another number on the timer. Report this number to me and write it down in your protocol. Then, again, there will be a 15 sec. pause, followed by the next trial, and so on.

When we have gone through the whole series of trials in this way, there will be a 10 min. break, followed by another series of comparisons, to complete each of the trials initiated before the break. During the break, the cards will be put back in the stands in their original order.

When the experiment is resumed after you have returned to the room following the break, I will begin by saying 'CARDS'. Then, as before, you are to turn over the first pair of cards of sets 1 and 2, compare them with regard to their numbers being odd or even, and place them face down on the table. A few seconds after that I will say the word 'DIGIT'. Then check the number you have earlier produced and recorded for this trial in your protocol, turn over the next card of set 3, and compare its number with the number you just read in the protocol. Then mark the result of the TWO comparisons as 'same' or 'different' under their respective headings in the protocol" (indicating) "and say 'READY'. Obviously, you must keep the first comparison - the one between the cards of sets 1 and 2 - in memory while performing the second- the one between the record of the number you produced earlier on the timer, and the one on the card of set 3. This completes the trial.

After a 15 sec. pause, the same procedure is repeated for the next trial: when I say 'CARDS', turn over the next pair of cards of sets 1 and 2 and compare their numbers. Then, when I say 'DIGIT', compare the recorded number with the one on the next set 3 card, score the result of the comparison - same or different - and say 'READY'.

This procedure is then repeated until all trials are completed. - Do you have any questions? - Then let us start up with two practice trials, using a 1/2-min. break instead of 10 min."

EUR.J.PARAPSYCHOL., V.5, 1984

CRITICISM AND CONTROVERSIES IN PARAPSYCHOLOGY - AN OVERVIEW

Eberhard Bauer
University of Freiburg

"When the main line of the history of parapsychological research is considered, as it tries to achieve a place among the recognized sciences, it immediately becomes apparent that the endless controversies and discussions associated with this enterprise which are still continuing to this day, are of a different nature than the usual conflicts in science. Apparently this debate touches upon more vital and essential values and issues than is normally found in polemics, for instance, about the acceptability or possible consequences of a technological finding or the introduction of a new chemical product. Especially the occasionally bitter tone of the discussions and the fact that they often become personal and violate common sense, is indicative of the inflammatory character of the issue" (Servadio, 1958, 1). This statement from the Italian psychoanalyst and parapsychologist Emilio Servadio highlights the peculiar position which these controversies about parapsychology and the critical attitudes towards its scientific nature occupy. There is no doubt that the problem of the existence of paranormal phenomena can be considered as one of the most controversial research topics in the history of science. It is even possible to view the history of parapsychology as the history of its controversies. Unlike in other scientific disciplines these controversies are not so much related to the interpretation of certain phenomena but refer to the very existence of the phenomena themselves. As will be demonstrated this is the reason why even the most competent judges do not agree about the essentials of paranormal research and reach different conclusions.

First of all even the question of competence is in dispute. Who is entitled to be considered as a 'parapsychologist', and vice versa, who is allowed to act as a critic in this area? It is not difficult to see that a homogeneous group of parapsychologists characterized by certain qualifications does not exist. The necessity of a curriculum preferably on an academic level and of professionalisation is well recognized (Shapin and Coly, 1976; Johnson, 1977); but without an organisational basis, financial support, and the corresponding acceptance by the scientific community, realisation is only possible on a limited scale. In short, there are no authorities in parapsychology in the sense of representatives of an accepted body of opinions, who are supported by most scientists involved. At best one can say that there are 'experts' although in this context the meaning of this term remains uncertain. In an instructive discussion about "areas of agreement between the parapsychologist and the skeptic", R.A. McConnell, himself an active psi-researcher, argues "unless you are willing and able to spend years training yourself in psychology, physics, and in the sociology of science, you cannot make a competent decision about the quality and conclusiveness of the experimental evidence for parapsychological phenomena" (McConnell, 1976, 304). Judged on such criteria the United States can perhaps count on "two dozen reasonably qualified and active research workers in parapsychology (McConnell, 1976, 308).

The same applies to the qualifications of the critics. The often applied dichotomy - parapsychologists believe in ESP, critics or skeptics do not - is simplistic as well as wrong. From many examples in the literature it can be demonstrated that parapsychologists are the most severe and competent critics of themselves and their own research. Just consider the names of Besterman, Dingwall, Hodgson, W.F. Prince, the contributions published in Murchison (1927), Angoff and Shapin (1971) and the discussion by John Beloff (1972, 1975) about the skeptic's position - just to name a few. The role of the self-appointed sceptic or professional debunker may seem prestigious in the public opinion (note 1) but often lacks factually based or logically acceptable argumentation (see for example Bender, 1964; Buchel, 1976; Bauer and von Lucadou, 1980).

Another controversy concerns the boundaries of the field. If the paradigms of the Rhinean School (Nilsson, 1975, 1976) which for forty years dominated research are accepted, then the only firmly settled parapsychological subject matter consists of extrasensory perception (ESP) and psychokinesis (PK). This explains why Rhine considered the "occult wave" (Bender, 1976, 7) which became prominent in the Western countries during the seventies and included acupuncture, Kirlian photography and astrology, as very dangerous for the image of parapsychology as an experimental science. Particularly because many parapsychologists appeared to take a positive attitude towards such pop-topics (to which can also be added ufology, Bermuda-triangle or pyramid-forces), Rhine warned that "Parapsychologists had better give some thought to the fact that their kind of psi is no longer nearly as securely under their own social control as in the past. The time has come when we who work with psi need to decide whether we really do know where we belong and just what our territory is. - - - Is there any other experimental science that rests on such a slight basis of uniformity and standardization?" (Rhine, 1972, 175).

If Schmeidler's questionnaire study (Schmeidler, 1971) can be considered as representative then it appears that the members of the Parapsychological Association at least concur that ESP is a proven phenomenon and that there is no reason to provide again and again new evidence (this may perhaps be too optimistic; see for example the recent enquiry among P.A. members by McConnel and Clark, 1980). But apart from that the opinions among leading parapsychologists are evenly divided as to which psi modality should be empirically studied. Rhine (1974b) for instance takes the position - criticized by Thouless (1973) - that a large number of parapsychological research topics, such as out-of-the-body experiences, the survival problem, retrocognition, psychometry and even telepathy are basically insolvable problems which cannot be studied empirically as it is impossible to eliminate clairvoyance as potential alternative hypothesis. It is safe to assume that this dilemma is not simply a semantic one. It reflects principally different theoretical models which have of course consequences for the empirical testability of the hypotheses derived therefrom. This picture of parapsychological diversity makes it rather easy for critics to compile from the literature a collection of widely varying statements and opinions (see for a recent example Alcock's coup de grace for parapsychology and Palmer's articulate rejoinder; Alcock, 1981; Palmer, 1983) which can be used to paint a livid picture of the most absurd consequences from research in this area, for instance, from a juridical point of view.

An overview of the history of a hundred years of research in parapsychology allows us to detect rather typical forms of pro and contra argumentation which influence the structure of the controversies in a remarkable way (see for instance the overviews given by Nicol, 1956; Crumbaugh, 1966; Dommeyer, 1966; Ransom, 1971). The address in 1882 by the first president of the British Society for Psychical Research (S.P.R.), Henry Sidgwick, is typical of how the pioneers of this research took it for granted how they would meet the objections of the scientific world. Sidgwick speaks of 'sufficient evidence', that is, "evidence that will convince the scientific world" (Sidgwick, 1882a, 9). Sidgwick elaborates on this in his second presidential address: "...if they will not yield to half-a-dozen decisive experiments by investigators of trained intelligence and hitherto unquestioned probity, let us try to give them half-a-dozen more recorded by other witnesses; if a dozen will not do, let us try to give them a score; if a score will not do, let us make up the tale to fifty" (Sidgwick, 1882b, 67). Thus the opposition should be gradually silenced and recognition of parapsychology enforced (note 2) by applying this principle of cumulative evidence, i.e. by adding more and more proof for the existence of ESP.

Closely related is the principle of reputable testimony: it became more or less standard procedure especially for sessions with physical mediums to involve large numbers of personalities with established reputation as observers in order to use their testimonies regarding the gennuineness of the phenomena in question to change the opinion of the scientific community. However, the controversy around the 'physical mediumship' which broke out between the two world wars primarily about the work of Schrenck-Notzing could not be resolved that way. The positions of both adherents and opponents remained basically irreconcilable (note 3).

When in the beginning of the thirties J.B. Rhine came forward with his experimental-statistical ESP research it seemed that this would create a change in the discussion. For the first time a number of independent researchers accepted a common methodology and terminology and applied it to a specified problem. It was also the first time that the scientific community was challenged by an excess of experimental results achieved under laboratory conditions by conventional methods and with unselected subjects.

The reaction of the scientific community to the proposed methodology of card experiments was accordingly animated. Between 1934 and 1940 about 60 critical publications by 40 authors appeared, mostly in the psychological literature, dealing with nearly every aspect of the experimental conditions and statistical evaluation (Honorton, 1975a). Most of the criticisms raised can be classified in three groups (Pope and Pratt, 1942). The first group concerns the mathematical-statistical assumptions of the evaluation techniques which were applied; the second the validity of the experimental procedures and the third the logic of the interpretation of the results in terms of the ESP hypothesis. The overview published in 1940 (Pratt et al, 1940) of all the main experimental research from 1882 till 1940 - the 'bible' of experimental parapsychology - lists and discusses 35 alternative hypotheses. To these belong erroneous statistical methods, improper selection of data, insufficient shuffling of target decks, optional stopping, unconsciously motivated errors in recording and checking target and response sequences, insufficiently eliminated sensory cues (unconscious whispering, marked cards) and finally incompetence and gullibility of the experimenters. Of the 142 publications from the previous 60 years only six turned out to be sufficiently robust to withstand all these objections, thus according to the authors providing valid evidence for paranormal cognition. These six are all experiments carried out in the Duke laboratory since 1927.

By applying such objective procedure Rhine and his collaborators to a large extent succeeded in silencing the main opposition by academic psychologists. Not in the least because they adapted their research in accordance with valid criticisms. Hence although the reality of ESP was not generally accepted, in the beginning of the forties at least agreement existed about what a proper ESP experiment should look like (Honorton, 1975b). When in 1943 the research program of the Duke laboratory expanded to include research in psychokinesis (influencing the throwing of dice) the criticisms remained restricted to those offered by the British parapsychologists who were mainly motivated by their lack of success in repeating these experiments. It was not before 1962 that the American psychologist Edward Girden published a fundamentally critical evaluation of 200 PK experiments and concluded that "evidence of PK as psychological phenomenon is totally lacking. And this deficiency will persist until the effect is produced in the presence of a specified psychological variable, and the effect does not appear in its absence" (Girden, 1962, 387).

Pratt (1964) objected that Girden exaggerated the defects of the experiments under consideration (for instance lack of strict experimental procedures, bad control of dice bias, improper evaluation of inhomogeneous data) and that he had ignored the experiments to which such objections were not applicable. Further information about the complexity of the problems relevant to the PK controversy and the differences between the opinions of the parapsychologists involved can be obtained by consulting the relevant literature (see Girden, Murphy, Beloff, Eisenbud, Flew, Rush, Schmeidler, Thouless, 1964).

After the successful completion of the "ESP controversy" in the sense that the opposition became silent at the end of the thirties Rhine took it for granted that only time was needed before parapsychology would be fully integrated in the psychological sciences (Nilsson, 1975, 1976). But this hope proved futile. In the next 15 years the 'establishment science' (Honorton, 1975b) took hardly notice of parapsychological research. The active confrontation failed to materialize. It was not before the middle of the fifties that the controversy erupted again. The immediate cause were two publications in the perhaps most influential interdisciplinary scientific journals 'Nature' and 'Science'. The Oxford logician G. Spencer Brown (1953) gave a new twist to the 'statistical controversy' in parapsychology by directing his criticism not against technical details of application of statistical procedures, but against the basic assumptions of probability theory itself. He disputed the common procedure in parapsychology to infer, from the 'improbability' of the result of the statistical evaluation the existence of ESP, despite the lack of repeatability and of demonstrable patterns in the phenomena. Such a naive way to infer a 'cause' from 'significance' was also criticized by the German mathematician Tornier. He argued that statistics is only a research tool and can never itself provide 'proof' (Tornier, 1959, 115) (note 4). Thornier's criticism was discussed at length in German parapsychology; the various positions concerning this controversy can be found in Bender (1959), Mischo (1974), Krengel and Liese (1978), and especially Timm (1979).

The most radical and perhaps most influential criticism so far directed against parapsychology up to that date was offered in an extensive discussion by the chemist George Price in his article 'Science and the supernatural', published as leading article in Science (Price, 1955). He started by admitting that the opposition of parapsychology has been practically silenced by an impressive number of careful experiments and intelligent argumentation. However, in view of the fact that the existence of ESP must be considered in conflict with current theories in science, Price was forced to conclude that all significant results in parapsychology which cannot be explained by faulty experimental procedures, statistical errors or unconscious use of sensory cues, had to be due to "deliberate fraud or mildly abnormal mental conditions" (Price, 1955, 360). Deliberate fraud by the investigator as an alternative for psi- under such a premise Price discussed a number of scenarios how fraud could have lead to the very significant results obtained by the British matematician Soal (Soal and Bateman, 1954) even though he did not provide factual evidence. Apart from the reactions of the scientists who were personally attacked (Rhine, 1956; Soal, 1956) especially Meehl and Scriven (1956) drew attention to two untenable presuppositions in Price's argumentation: Firstly that ESP is in conflict with modern science and secondly that modern science in its present shape should be correct and complete. In any case, seven years later Price withdrew his suspicion of Rhine and Soal as frauds as 'highly unfair' (Price, 1972, 356) (note 5). Nevertheless both arguments, the a priori improbability of ESP and the possibility of fraud on the part of the experimenter, were taken up and extended in the book by the British psychologist C.E.M. Hansel, published in 1966: 'ESP, a scientific evaluation'. The non-parapsychological world seemed to consider this book as the final word to be wasted on the subject (see Slater, 1968). According to Hansel the process investigated in parapsychology is: "both hypothetical and a priori extremely unlikely" (Hansel, 1966, 17). Any possible known cause of the results, including conspiracy by the participants of the experiment to cheating, is far more likely to be responsible for it than the hypothetical process (ESP) under consideration.

In the analysis of four experiments, of which three belong to the 'classical' conclusive ESP experiments: the Pearce-Pratt series and the Pratt-Woodruff experiment both from the early Duke period and Soal's experiments with Mrs Stewart and Basil Shackleton as well as the Soal-Bowden experiments with three Welsh schoolchildren, Hansel demonstrates with remarkable ingenuity how fraud could have been committed. According to Hansel this is sufficient to question any positive claim for convincing evidence of ESP (Hansel, 1966, 241). It is hardly possible to counter such charges of fraud, at least not as long as independent confirmation for the findings are lacking. But in the case of parapsychology the argument of fraud is not more plausible than in the case of other scientists (see for instance McConnell, 1975). The scandal around Rhine's successor, W.J. Levy, which erupted in 1974 and which resulted in world-wide news comments, demonstrates particularly the essential point that Levy's fraud was detected by his colleagues and that Rhine himself made it public (Rhine, 1974c) (note 6).

Hansel's critical approach to parapsychology was heavily criticised by pointing out the apparent bias of his arguments and on account of many factual errors and inaccuracies which makes it doubtful whether this work can be called a 'scientific' evaluation of psi (note 7). Nevertheless Hansel's penetrating criticism highlights a number of fundamental problems. The opinions regarding the importance of these problems differ in the parapsychological community.

In the first place one can consider in this connection the problem of repeatability (see especially Crumbaugh, 1966; and - more recently - the thorough discussion by Hovelmann, 1983). At a minimum it can be said that everybody agrees that parapsychology knows repeatable experiments but not repeatable results. According to Beloff (1972) an experiment with repeatable results can be considered as the description of an experimental procedure which, when applied by competent researchers, "must work at least 50 percent of time and, even more important, must not depend on the availability of a particular individual as subject" (Beloff, 1972, 198). But the opinions under parapsychologists about this matter varies. For instance Beloff in agreement with Crumbaugh (1966, 526) and Dommeyer (1966) concludes that parapsychological results will only be generally accepted by the scientific community when at least one repeatable effect can be demonstrated. Proposals to modify the concept of repeatability, along the lines suggested by LeShan (1966) or Murphy (1971), make it dependent on the specific character of the research object. For parapsychology this would imply that an 'intrasubjective' repeatability exists in the sense, that over several years a subject achieves positive results with different investigators. An example is the 'focussing effect' of Pavel Stepanek (Keil, 1977). The sort of 'internal repeatability' found in experiments of the Maimonides group with telepathic dream induction (Ullman et al, 1973) or in the phenomenology of paranormal metal bending (Hasted, 1977) could also be rated as such. At any rate, the demand for repeatability remains a fundamental methodological problem in parapsychology. But this holds not only for parapsychology but also for the behavioral sciences. For instance, in psychology the results are similarly characterised by widespread inconsistencies, by non-repeatability and non-predictability (Maschewsky 1977, 212). Regarding repeatability Honorton (1975b) feels that compared to certain fields in psychology parapsychology is even in a better state.

The attitude towards the repeatability issue and its epistemological foundations has far reaching consequences for a number of related 'subproblems' which can only be shortly mentioned here. For instance the problem of selective reporting of only positive results could lead to a distorted picture of the actual achievements of research (for this problem see the discussion between Rhine, 1975 and Beloff et al, 1976). Another related problem is the empirical verification of hypotheses. By combining in an uncritical way different 'effects', like psi-missing decline-effects or influence of experimenter bias, it becomes in principle possible to interpret each outcome of a parapsychological experiment in support of the psi hypothesis. The danger of such a strategy which ensures the immunity of the psi interpretation against nearly all criticism is reinforced by the generally applied terminology in parapsychology. For instance it is asserted that a certain phenomenon can be 'explained by ESP' (see Mundle, 1971, 20). Such an expression neglects that the ESP concept has no explanatory power but should be considered merely a verbal convention to label a certain as yet unexplained group of phenomena (more about this in Staub, 1978). The frequently discussed observation that psi often fails to appear when skeptical observers (for instance magicians) or researchers are involved can be interpreted, from a psychological point of view, as an indication for the dependency of psi phenomena on complex psychological conditions, a delicate affective field (Bender, 1976) between the participants in a psi experiment. In other words, there may be quite a number of unknown conditions which requires new strategies for dealing with.

A further problem for the controversial status of parapsychology, indirectly related to the issue of repeatability but of exceeding importance, concerns the remarkable erosion of evidence. This 'evaporative effect' as Scriven called it (quoted by Eisenbud, 1963, 251) means that some initially so convincing results of parapsychological research seem to lose their strength with later reevaluations. When time passes even the researcher will eventually become affected by the destructive influence of doubt. John Beloff (1972) as president of the Parapsychological Association gave a lively illustration of this 'genesis of doubt' (Rogo, 1977) with examples of prominent parapsychologists (note 8). The 'will to believe' of parapsychologists as assumed by the skeptics seems more like a 'will to disbelieve' their own experiments and observations. This "principle of retroactive dissonance" (B. Inglis) can be nicely demonstrated by for instance the famous S.P.R. investigations of Eusapio Palladino in 1908 (see Rogo, 1977). The 'erosion of evidence' is one of the most stable traits in the history of parapsychology. Nearly every 'classical case', every 'conclusive experiment' has been subjected to this 'test of time', the process of re-evaluation based on new evidence and re-interpretation viewed from a different perspective. This reconstruction of evidential material often accompanied with much controversy can only be touched upon here; it fills virtually thousands of pages as the history of psychical research (see for example the Proceedings of the SPR) makes it abundantly clear (see especially Inglis, 1977). The disputes about the genuiness of William Crookes' experiments with D.D. Home and the 'materialisation' medium Florence Cook last already more than a hundred years (Medhurst and Goldney, 1964; Medhurst, 1972). Trevor H. Hall (1962) for instance tries to prove in his much debated book 'The Spiritualists' that Crookes was Florence's lover and helped her to cheat during sessions (see the discussion between Stevenson, 1963, 1964 and Hall, 1964a). Especially Hall's investigations who like a detective tries to detect 'weak spots' in the old S.P.R. experiments (see Campbell and Hall, 1968) give constantly fresh impetus to the historically oriented controversies, as in the case of Hall's book on one of the founders of the S.P.R., Edmund Gurney. According to Hall (1964b)

Gurney withheld indications for fraud and later committed suicide (for a detailed critical appraisal see Nicol, 1966, and Hall's rejoinder, 1968). So the controversies continue (for a recent example see Brandon, 1983). The special studies of famous cases, like Tietze's (1973) study of Margery, Rogo's (1975) study on Palladino or Anita Gregory's (1977) study of Rudi Schneider, demonstrates the typical pattern of the scientific controversies in parapsychology. To this pattern belongs the emotional polarisation of the antagonists, the competence claims, the committees to evaluate the 'conclusive evidence', the offering of awards, etc..

The 'erosion of evidence' affects also those experiments which for a long time were considered as the most solid data of parapsychology. The tragic irony of the famous Soal-Shackleton series of 1941-1943 on precognitive telepathy (Soal and Bateman, 1954) with its experimental design aimed at eliminating possible experimenter fraud but to which Soal himself gave rise to suspicion (the only afterwards admitted loss of the original protocols; use of the random tables differs from as reported; the allegation by Grete Albert that she saw Soal changing figures; see for this controversy Scott et al, 1974; Scott and Haskell, 1975; Markwick, 1978) (note 9). In the same vein for already decennia the Pratt-Woodruff experiment constitutes a platform for accusations of fraud by critics (Pratt, 1976). These examples teach us at least one lesson. The conclusive experiment convincing every sceptic of the the existence of psi does not exist. It is an illusion to make the break-through to scientific recognition depending on the 'perfect' experiment. In his review of Hansel's book Stevenson wrote the remarkable sentence: "If we give up the idea of a fraud-proof experiment we ought also to give up the idea that our experiments are in any way conclusive or can be regarded as proof" (Stevenson, 1967, 263f). He rather argues in favor of some agreed-upon standards developed in cooperation between researchers in parapsychology and critics inside and outside the field for the evaluation of a specific experiment (Stevenson and Roll, 1966). The apparent impossibility of the 'decisive experiment' is confirmed by Nicol's observation that even "psychical researchers of undoubted authority do not agree among themselves as to whether some of the leading experiments are conclusive evidence for paranormality" (Nicol, 1956, 29).

Considering such a situation Coover's 'fagot-theory' (Coover, 1927, 233) offers some perspective. Although each piece of evidence, each branch so-to-say, can be criticised and in principal be refuted, together they constitute a strong bundle of evidence. On the other hand it is equally possible to defend a 'chain' model (Beloff, 1976, 93). The chain of evidence for ESP is as strong as its weakest link. It seems therefore unavoidable that parapsychologists often apply subjective criteria in weighing the evidence. For Rhine (1974a, 113), for instance, the unexpected post hoc discoveries of 'fingerprints of psi' in the card guessing and dice experiments, the decline effects and U-curves, constitutes convincing evidence. Surely much misunderstanding would be avoided if subjective evidence could be kept strictly separated from compelling scientific evidence, although it is questionable especially in the case of parapsychology whether that is possible. Is not personal motivation, the experimenter effect and a positive attitute towards psi an essential condition for eliciting psi?

One group of critics considers the answer to this question as the very solution of the mystery of psi. According to this argumentation first offered by Moll (1929), further developed by Gubisch (1961) and taken up by Prokop and Wimmer (1976) the gullible parapsychologists live in a joyful anticipation of the occult and cover their superstitions with a pseudo-scientific cloak. Thus the whole field of parapsychology only exists because of the perhaps psychologically abnormal motivation of parapsychologists. Especially the German critic Wilhelm Gubisch reduces the whole problem of ESP to the 'psychological structure of the believers in the occult' (Gubisch, 1961, 98). As a pseudo-clairvoyant in his 'experimental demonstration of ESP' he collected from the general public valuable material about the gullibility and the will to believe. But this very example can also be used to demonstrate how Gubisch's motivation as a debunker, analysed according to the principles of research in social perception, distorted the way of handling his data (see for examples Neuhausler, 1964). But despite the intensity of his negative attitude at least Gubisch demonstrates his awareness of the possible consequences of the paranormal (Bender, 1964). With others the problem is reduced to one of a purely psychological nature (see for instance Wimmer, 1973). Already W.F. Prince (1930) observed that even when scientifically educated persons enter the field parapsychology and pass the 'enchanted boundary' they suddenly appear to become one-sided in the information they collect and to ignore arguments. In short, they react so irrational in their opposition as would be unthinkable inside their own field. Apparently firmly rooted defenses against the acceptance of the paranormal lie behind the rational discussions. Servadio (1958) when interpreting this defense proposes a psychodynamically based 'disbelieve reaction' to parapsychological phenomena. In Eisenbud's (1963; 1966) speculations the defense against psi is part of nature itself, and even parapsychologists are prevented from gaining experimental control over these powers by an 'unconscious sabotage' directed against their own efforts. LeShan (1966) applies Festinger's model of cognitive dissonance. The psychological motivation to reject paranormal phenomena originates in their very observation, which is in conflict with the familiar social-cultural context and thus creates a threatening conflict. An explanation is attempted here by means of depth and social psychological concepts, namely that neither the amount nor the scientific quality of evidence for parapsychological phenomena contributes to its social acceptance. This becomes even more clear when considered from the point of view of the history and sociology of science. Here the controversial status of parapsychological research becomes the prime example of the general problem in the development of science, i.e. that the acceptance of new phenomena and theories is hardly influenced by the objective state of evidence (Ferrera, 1977). Especially the sociological study of the parapsychological community can serve to demonstrate the close association between the social organization of an innovative group and the reaction of the established sciences (Allison, 1973).

Among parapsychologists McConnell (1966) has been the first to interpret the controversial situation of parapsychology in terms of Thomas Kuhn's (1962) influential model for the development of sciences. In this model parapsychological data become 'anomalies' which are in conflict with the currently dominating 'paradigms' of the natural and social sciences and consequently provoke opposition. The picture of modern parapsychology - uncoordinated and random observations, conflicting experimental results, the lack of well-defined concepts, of generally accepted working hypotheses and related theories, the desintegration into competing schools, the emotionality of the controversies - are also the features of an immature preparadigmatic phase of science awaiting an 'Einstein of parapsychology' (Pratt, 1974) to guide the field into the realm of accepted sciences. To what extent this interpretation in terms of Kuhn's model is perhaps too optimistic or even misleading, is in itself the subject of current discussion (Shapin and Coly, 1977). But undoubtedly such 'metaperspectives' are of great value to determine the position of a 'protoscience'. Parapsychology as a study object for a 'relativistic sociology of science' (Rao, 1977) demonstrates the extent to which the scientific acceptability depends on the concensus of a group relying upon changing historical criteria. This further relieves parapsychology from concentrating on the existential question of the 'yes' or 'no' of her phenomena, of the concept of psi, again and again newly discovered and accepted by one group of people and then rejected and buried by others. In short, in this perspective the conflict around parapsychology becomes the touchstone for hidden and anthropological assumptions in our scientific worldview and research methodology. This too constitutes a challenge of parapsychology.

POSTSCRIPT

This slightly revised overview was written in 1977 and first published in German in the 15 volume "Kindler's Psychologie des 20. Jahrhunderts". It was primarily intended for psychologists and other educated people, who can be assumed te be rather ignorant about parapsychology and who probably have never heard of the 'European Journal of Parapsychology' and other professional journals in our field. Because of lack of space the topic had to be discussed in relatively few pages and therefore important aspects of the subject of the paper sometimes could only be touched upon.

One thing is for certain, even in 1984: The psi controversy is still with us. However, it appears that in the last years a start has been made towards a more rational and fair dialogue between proponents and critics of the paranormal (among the latter various representatives of CSISCOP). K.R. Rao, for instance, organized as part of the PA Convention in 1981 a symposium entitled 'Parapsychology and its critics: Implications for philosophy and sociology of science', in which a number of recently published critical views on parapsychology (Girden, Diaconis, Moss and Butler, Gibson, Kurtz) were discussed. Similarly at the Centenary-Jubilee Conference (SPR 1882-1982 and PA 1957-1982) a symposium was held, entitled "The case for skepticism", in which among others C. Scott, S. Blackmore, P.H. Hoebens and R. Hyman took part. The 'Zetetic Scholar', edited by sociologist M. Truzzi has developed in the last few years into one of the best sources of information about criticism of research in parapsychology, with contributions from insiders and outsiders. Prominent are the 'Major Dialogues' between parapsychologists and skeptics (see especially Hyman, Beloff, Westrum, Hovelmann). In addition, a number of books taking a critical view of parapsychology have recently been published, among them the second edition of Hansel's book but also books by Alcock, Marks and Kammann, and Abell and Singer, which evoked several extensive evaluations from the parapsychological community.

NOTES

1. Symptomatic for this is the behavior of some members of the Committee for the Scientific Investigation of Claims of the Paranormal (CSICOP) founded in 1976 with which the best-known critics of parapsychology like Gardner, Hansel, Hyman and Randi are associated. Before issuing their own journal 'The Sceptical Inquirer' the 'Humanist' published by the philosopher Paul Kurtz was the mouthpiece of the CSICOP. The pretention of a rational evaluation of paranormal phenomena and the applied methodology has evoked some sharp criticisms by parapsychologists, see for instance Rockwell et al (1978) and Kurz et al (1978).

2. From the point of view of the sociology of sciences it certainly would be rewarding to compare the various presidential addresses of the S.P.R. in order to study the development in history of what is considered as 'established parapsychological knowledge', how that knowledge was acquired and the progress parapsychology has made regarding scientific recognition.

3. The controversy is most clearly presented in the 'Drei-Manner Buch' (three-men book) of Gulat-Wellenburg, v. Klinckowstroem, Rosenbusch (1925), the 'Sieben-Manner Buch' (seven-men book) published by Schrenck-Notzing (1926), and the subsequent discussions in the 'Zeitschrift fur Parapsychologie' and the 'Zeitschrift fur Kritischen Okkultismus'. An evaluation of the opposite views is presented by the Swiss psychiatrist Bleuler (1930).

4. Tornier expanded his criticism into 'Rhine - Fall of the parapsychologists', which was met with approval from critics like Prokop and Wimmer (1976, 122). The decisive mathematical rebuttal was, not considering Buchel (1975, 170), provided by Krengel and Liese (1978) and especially by Trimm (1979).

5. In an addition to the reprint of his article by French (1975, 373) Price states: "that I have myself become guilty of accepting and trying to follow (in a rather radical way) that strange system of beliefs that I accused Rhine and Soal of trying to promote, and consequently I now believe in much worse things than ESP".

6. The fundamental importance of the fraud and deceit argumentation is discussed by Muller (1980).

7. See for instance the critical evaluations by Honorton (1967), Stevenson (1967) and Medhurst (1968). Especially instructive are the positions of Eysenck, West, Beloff, Stevenson, the review by Slater (1968) and the discussion between Hansel and Slater (British Journal of Psychiatry, 114, 1968, 1471-1480; and ibid 115, 1968, 743-745).

8. Compare also the resigned attitude of W. James in his 'Final Impressions of a Psychical Researcher', of 1909, reprinted in Murphy and Ballou, 1969, 309-325, especially page 310.

9. Meanwhile the astute analysis by Betty Marwick (1978) leaves little doubt that Soal manipulated the target sequences of the Shackleton experiment. The motive for Soal's behavior remains unclear. However, undoubtedly experimental parapsychology lost an important piece of evidence and the adherents of the 'Psi=fraud' thesis scored another point.

REFERENCES

Alcock, J.E. 'Parapsychology: science or magic? A psychological perspective', New York, Pergamon, 1981.

Allison, P.D. 'Social aspects of scientific innovation: The case of parapsychology'. M.A. thesis University of Wisconsin, 1973.

Angoff, A and Shapin, B. 'A century of psychical research: The continuing doubts and affirmations'. New York, Parapsychology Foundation, 1971.

Bauer, E. and von Lucadou, W. 'Parapsychologische Forschung und wissenschaftliche Methodik', Z. Parapsychol. Grenzgeb. Psychol., 1980, 22, 51-70.

Beloff, J. 'Parapsychology as science'. International J. Parapsychology, 1967, 9 (No.2), 91-97.

Beloff, J. 'Belief and doubt'. Research in Parapsychology, 1972, 1, 189-200.

Beloff, J. 'Psychological sciences: A review of modern psychology'. London, Crosby Lockwood Staples, 1975, 282-320.

Beloff, J., Broughton, R.S. and Millar, B. 'Letter to the editors'. J.o.P., 1976, 40, 88-91.

Bender, B. 'Parapsychologie und Rechtsordnung'. Neue Jurist. Wschr., 1977, 24, 1089-1091.

Bender, H. 'Editorial'. Z. Parapsychol. Grensgeb. Psychol., 1959, 3, 81-88.

Bender, H. 'Formen der Einstellung zur Parapsychologie', Z. Parapsychol. Grenzgeb. Psychol., 1964, 7, 85-92.

Bender, H. 'Parapsychologie, ihre Ergebnisse und Probleme', Frankfurt/M., Fischer, 1976.

Bleuler, E. 'Vom Okkultismus und seinen Kritiken', Z. Parapsychol. Grenzgeb. Psychol., 1930, 5, 654-680.

Brandon, R. 'The spiritualists. The passion for the occult in the nineteenth and twentieth centuries', London, Weidenfeld and Nicholson, 1983.

Brown, G.S. 'Statistical significance in psychical research', Nature, 1953, 172, 154-156.

Buchel, W. 'Zur Kritik an der Parapsychologie', Z. Parapsychol. Grenzgeb. Psychol., 1976, 18, 161-186.

Campbell, J.L. and Hall, T.H. 'Strange things', London, Routledge and Kegan Paul, 1968.

Coover, J.E. 'Metapsychics and the incredulity of psychologists', In C. Murchison 'The case for and against psychical belief', Worcester, Clark University, 1927, 229-264 (New edition New York, Arno Press, 1975).

Crumbaugh, J.C. 'A scientific critique of parapsychology', Intern. J. Neuropsychiat., 1966, 2, 423-531.

Diaconis, P. 'Statistical problems in ESP research', Science, 1978, 201, 131-136.

Dommeyer, F.C. 'Parapsychology: old delusion or new science', Intern. J. Neuropsychiat., 1966, 2, 539-555.

Eisenbud, J. 'Psi and the nature of things', International J. Parapsychol., 1963, 5, 245-273.

Eisenbud, J. 'The problem of resistance to psi', In Roll, W.G., Proceedings of the Parapsychological Association, 1966, 3, 63-79.

Ferrera, L. 'Beitrage der Wissenschaftstheorie zu umstrittenen Forschungsgebieten', In I. Brand (ed.) 'Unerklarliche Himmelserscheinungen aus alterer und neuerer Zeit', Munchen, Mufon-Ces, 1977, 31-56.

French, P.A. 'Philosophers in wonderland. Philosophy and psychical research', Saint Paul, Llewellyn, 1975.

Girden, E. 'A review of psychokinesis', Psychological Bulletin, 1962, 59, 353-388.

Girden, E., Murphy, G., Beloff, J., Eisenbud, J., Flew, A., Rush, J.H., Schmeidler, G. and Thouless, R.H. 'A discussion of psychokinesis', Intern. J. Parapsychol., 1964, 6, 26-137.

Gregory, A. 'Anatomy of a fraud: Harry Price and the medium Rudi Schneider', Annals of Science, 1977, 34, 499-549.

Gubisch, W. 'Hellseher, Scharlatane, Demagogen? Kritik an der Parapsychologie', Munchen, Basel, Reinhard, 1961.

Gulat-Wellenberg, W.v., Klinckowstroem, C.v. and Rosenbusch, H. 'Der physikalische Mediumismus', In M. Dessoir (ed) 'Der Okkultismus in Urkunden', Part I, Berlin, Ullstein, 1925.

Hall, T.H. 'The spiritualists: The story of Florence Cook and William Crookes', London, Duckworth, 1962.

Hall, T.H. 'Reply to Dr. Ian Stevenson's review of "The Spiritualists", J. A.S.P.R., 1964, 58, 57-65. (a)

Hall, T.H. 'The strange case of Edmund Gurney', London, Duckworth, 1964. (b)

Hall, T.H. 'Some comments on Mr. Fraser Nicol's review', Intern. J. Parapsychol., 1968, 10, 149-164.

Hansel, C.E.M. 'ESP: A scientific evaluation', New York, Scribner's, 1966.

Hasted, J.B. 'Physical aspects of paranormal metal bending', J. S.P.R., 1977, 49, 583-607.

Honorton, C. Book review (of Hansel, 1966), J.o.P., 1967, 31, 76-82.

Honorton, C. 'Error Some Place', J. Communication, 1975, 25, 103-116. (a)

Honorton, C. 'Has science developed the competence to confront claims of the paranormal', Research in Parapsychology, 1975, 3, 199-223. (b)

Hovelman, G.H. 'Zum Problem der Wiederholbarkeit parapsychologischer Experimente', Z. Parapsychol. Grenzgeb. Psychol., 1983, 25, 29-54.

Inglis, B. 'Natural and supernatural. A history of the paranormal from earliest times to 1914', London, Hodder and Stoughton, 1977.

James, W. 'The final impressions of a psychical researcher', In G. Murphy and R.O. Ballou (eds) 'William James on Psychical Research', New York, Viking, 1969, 309-325.

Johnson, M.U. 'Who is a parapsychologist?', E.J.P., 1977, 2, 1-3.

Keil, H.H.J. 'Pavel Stepanek und der Fokussierungs-Effekt', Z. Parapsychol. Grenzgeb. Psychol., 1977, 19, 1-22.

Krengel, U. and Liese, M. 'Begrundung und Kritik der Statistischen Auswertung parapsychologischer Experimente', Z. Parapsychol. Grenzgeb. Psychol., 1978, 20, 185-211.

Kuhn, T.S. 'The structure of scientific revolutions', Chicago, Univ. of Chicago Press, 1962.

Kurtz, P., Rockwell, T., Rockwell, W.T. 'The humanist's crusade against parapsychology: A discussion', J. A.S.P.R., 1978, 72, 349-364.

LeShan, L. 'Parapsychology and the concept of the "repeatable experiment"', Intern. J. Parapsychol., 1966, 8, 133-145. (a)

LeShan, L. 'Some psychological hypotheses on the non-acceptance of parapsychology as a science', Intern. J. Parapsychol., 1966, 8, 367-385. (b)

Markwick, B. 'The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation', Proc. S.P.R., 1978, 56, 250-277.

Maschewsky, W. 'Das Experiment in der Psychologie', Frankfurt/M, Campus, 1977.

McConnell, R.A. 'ESP research at three levels of method', J.o.P., 1966, 30, 195-207.

McConnell, R.A. 'The motivations of parapsychologists and other scientists', J. A.S.P.R., 1975, 96, 273-280.

McConnell, R.A. 'Areas of agreement between the parapsychologist and the skeptic', J. A.S.P.R., 1976, 70, 303-308.

McConnell, R.A. and Clark, T.K. 'Training, belief, and mental conflict within the Parapsychological Association', J.o.P., 1980, 44, 245-268.

Medhurst, R.G. 'The fraudulent experimenter: Professor Hansel's case against psychical research', J. S.P.R., 1968, 44, 217-232.

Medhurst, R.G. 'Crookes and the spirit world', New York, Taplinger, 1972.

Medhurst, R.G. and Goldney, K.M. 'William Crookes and the physical phenomena of mediumship', Proc. S.P.R., 1964, 54, 25-157.

Meehl, P.E. and Scriven, M. 'Compatability of science and ESP', Science, 1956, 123, 14-15.

Mischo, J. 'Die Hypothese des Entlarvungstechnikers', Z. Parapsychol. Grenzgeb. Psychol., 1964, 7, 126-145.

Mischo, J. 'Parapsychische Phanomene im quantitativ-statistischen Modell. Teil II: Thesen der Kritiker', Z. Parapsychol. Grenzgeb. Psychol., 1974, 16, 125-147.

Moll, A. 'Psychologie und Charakterologie der Okkultisten', Stuttgart, Enke, 1929.

Mundle, C.W.K. 'Confusion about ESP" and skepticism about "ESP"'. In A. Angoff and B. Shapin (eds) 'A century of psychical research: The continuing doubts and affirmations', New York, Parapsychology Foundation, 1971, 17-30.

Murchison, C. 'The case for and against psychical belief', Worcester, Clark University, 1927.

Murphy, G. 'Report on a paper by Edward Girden on psychokinesis', Psychological Bulletin, 1962, 59, 520-528.

Murphy, G. 'The problem of repeatability in psychical research', J. A.S.P.R., 1971, 65, 3-16.

Neuhausler, A. 'Die Patentlosung', Z. Parapsychol. Grenzgeb. Psychol., 1964, 7, 93-125.

Nicol, J.F. 'Some difficulties in the way of scientific recognition of ESP'. In G.E.W. Wolstenholme and E.C.P. Millar (eds) 'Extrasensory Perception: A Ciba foundation symposium', London, Churchill, 1956, 24-37.

Nicol, J.F. 'The silences of Mr. Trevor Hall', Intern. J.o.P., 1966, 8, 5-59.

Nilsson, I. 'The paradigm of the Rhinean school', E.J.P., 1, 1, 45-60.

Nilsson, I. 'The paradigm of the Rhinean school. Part II., E.J.P., 1, 2, 45-57.

Palmer, J. 'In defense of parapsychology: A reply to James E. Alcock', Zetetic Scholar, 1983, 11, 39-70.

Pope, D. and Pratt, J.G. 'The ESP controversy', J.o.P., 1942, 6, 175-189.

Pratt, J.G. 'The Girden-Murphy papers on PK', J.o.P., 1963, 27, 199-209.

Pratt, J.G. 'Some notes for the future Einstein of parapsychology', J. A.S.P.R., 1974, 68, 133-155.

Pratt, J.G. 'New evidence supporting the ESP interpretation of the Pratt-Woodruff experiment', J.o.P., 1976, 40, 217-227.

Pratt, J.G., Rhine, J.B., Smith, B.M., Stuart, C.E. and Greenwood, J.A. 'Extra-sensory perception after sixty years', Boston, Bruce Humphries, 1940.

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

Price, G.R. 'Apology to Rhine and Soal', (Letter), Science, 1972, 175, 359.

Prince, W.F. 'The enchanted boundary. Being a survey of negative reactions to claims of psychic phenomena, 1820-1930', Boston, Boston Society for Psychical Research, 1930. (new edition: New York, Arno Press, 1975).

Prokop, O. and Wimmer, W. 'Der moderne Okkultismus. Parapsychologie und Paramedizin', Stuttgart, Fischer, 1976.

Ransom, C. 'Recent criticisms of parapsychology', J. A.S.P.R., 1971, 65, 289-307.

Rao, K.R. 'Some frustrations and challenges in parapsychology', J.o.P., 1977, 41, 119-135.

Rhine, J.B. 'Comments on "Science and the supernatural"', Science, 1956, 123, 11-14.

Rhine, J.B. 'Comments: Is parapsychology losing its way?', J.o.P., 1972, 36, 170-176.

Rhine, J.B. 'Security versus deception in parapsychology', J.o.P., 1974, 38, 99-121. (a)

Rhine, J.B. 'Telepathy and other untestable hypotheses', J.o.P., 1974, 38, 137-153. (b)

Rhine, J.B. 'A new case of experimenter unreliability', J.o.P., 1974, 38, 215-225. (c)

Rhine, J.B. 'Publication policy regarding nonsignificant results', J.o.P., 1975, 39, 135-142.

Rockwell, T., Rockwell, R. and Rockwell, W.T. 'Irrational rationalists: a critique of "The Humanist's" crusade against parapsychology', J. A.S.P.R., 1978, 72, 23-34.

Rogo, S.D. 'Eusapia Paladino and the structure of scientific controversy', Parapsychology Review, 1975, 6 (No.2), 23-27.

Rogo, S.D. 'Parapsychology and the genesis of doubt', Parapsychology Review, 1977, 8 (No.6), 20-22.

Schmeidler, G. ''Parapsychologists' opinions about parapsychology', 1971, J.o.P, 1971, 35, 208-218.

Schrenck-Notzing, A.v. 'Die physikalischen Phanomene der grossen Medien. Eine Abwehr', Stuttgart, Union, 1926.

Scott, C., Haskell, P., Goldney, K.M., Mundle, C.W.K., Thouless, R.H., Beloff, J., Pratt, J.G., Barrington, M.R., Stevenson, I. and Smythies, J.R. 'The Soal-Goldney experiments with Basil Shackleton: a discussion', Proceedings S.P.R., 1974, 56, 41-131.

Scott, C. and Haskell, P. 'Fraud in the Shackleton experiments: a reply to critics', J. A.S.P.R., 1975, 48, 220-226.

Servadio, E. 'Parapsychologie und "Unglaubigkeitsreaktion"', Z. Parapsychol. Grenzgeb. Psychol., 1958, 2, 1-9.

Shapin, B. and Coly, L. 'Education in parapsychology', New York, Parapsychology Foundation, 1976.

Shapin, B. and Coly, L. 'The philosophy of parapsychology', New York, Parapsychology Foundation, 1977.

Sidgwick, H. 'First presidential address', Proceedings S.P.R., 1882, 1, 7-12. (a)

Sidgwick, H. 'Address to the S.P.R., second meeting', Proceedings S.P.R., 1882, 1, 65-69. (b)

Slater, E. 'Book review (of Hansel, 1966)', British J. of Psychiatry, 1968, 114, 653-658.

Soal, S.G. 'On "Science and the supernatural"', Science, 1956, 123, 9-11.

Soal, S.G. and Bateman, F. 'Modern experiments in telepathy', London, Faber and Faber, 1954.

Staub, W. 'Beschreibungs- und Beobachtungssprache in der Parapsychologie - eine methodologische Ktitik', Master thesis, University of Freiburg, 1978.

Stevenson, I. 'Reflections on Mr.Trevor Hall's "The Spiritualists"', J. A.S.P.R., 1963, 57, 215-226.

Stevenson, I. 'Rejoinder to Mr.Trevor Hall's reply', J. A.S.P.R., 1964, 128-133.

Stevenson, I. 'An antagonist's view of parapsychology: a review of Professor Hansel's "ESP: A Scientific Evaluation", J. A.S.P.R., 1967, 61, 254-267.

Stevenson, I. and Roll, W.G. 'Criticism of parapsychology: an informal statement of some guiding principles', J. A.S.P.R., 1966, 60, 347-356.

Thouless, R.H. 'Summary of dialogue' (in News and Comments), J.o.P., 1973, 37, 69-70.

Tietze, T.R. 'Margery', New York, Harper and Row 1973.

Timm, U. 'Tornier und die Parapsychologie', Z. Parapsychol. Grenzgeb. Psychol., 1979, 21, 99-120.

Tornier, E. 'Die Arbeitshypothese "Antizufallswahrscheinlichkeit"', Z. Parapsychol. Grenzgeb. Psychol., 1959, 3, 90-119.

Tornier, E. 'Der "Rhine-Fall" der Parapsychologen', privately printed by the author, 1974.

Ullman, M., Krippner, S. and Vaughan, A. 'Dream Telepathy', London, Thurstone, 1973.

West, D.J. 'Reasons for the continuing doubt about the existence of psychic phenomena'. In A. Angoff and B. Shapin (eds) A century of psychical research: The continuing doubts and affirmations', New York, Parapsychology Foundation, 1971, 155-159.

Wimmer, W. 'Eine andere Wirklichkeit? Von Unfug der Parapsychologie', Deutsches Arzteblatt, 1973, 71, 732-739.

Wimmer, W. 'Parapsychologen als Sachverstandige?', Neue Jurist. Wschr., 1976, 1131-1133.

Eberhard Bauer
University of Freiburg
Department of Psychology
Belfortstrasse 16
D-7800 Freiburg
West Germany

EUR.J.PARAPSYCHOL., V.5, 1984

PSI TRANSMISSION OF ANGER AND ITS ENHANCEMENT THROUGH MEANING? METHODOLOGICAL AND STATISTICAL CONCERNS

Rex G. Stanford
St.John's University

Ephraim I. Schechter
Psychophysical Research Laboratories

Shulamith and Hans Kreitler have, in Volume 4, Number 2 of this journal (pp.119-241) reported a methodologically inovative study of an area which potentially has considerable interest for this field. The claims made in the title of the paper and, of course, elsewhere in it are made boldly and without any indication that the authors feel that further research is needed before firm conclusions can be made with any realism. We recognize the importance of these claims, should they be true. In our judgment, however, the Kreitlers' report fails to substantiate them. Our critique will focus on information that has not been provided but that we feel is necessary to interpret the results of this first study, and on reasons why we feel that the design and the statistical analyses of this study do not adequately test the Kreitlers' hypotheses. We have tried to describe the problems we see in ways that suggest improvements for future studies.

GENERAL PERSPECTIVE

The Kreitlers' study is sufficiently complex and its problems sufficiently great that a general perspective on the study will aid in understanding these problems.

Most ESP studies use as the dependent variable true ESP scores, but in the Kreitlers' study the scores assigned to individual subjects cannot, by the nature of things, be considered as ESP scores. (For example, one cannot specify mean chance expectation for them.) In such a study the occurrence of ESP can be inferred only through between-groups comparisons of a dependent measure (especially, here, emotionality). Specifically, all inferences about ESP must, in such a study, be based upon the influences of 'agent variables' (e.g., the agent - no agent contrast) upon the dependent measures taken from sessions with the receivers. Although the agent variable had four levels in the Kreitlers' study, only the agent - no agent contrast was important in terms of being statistically significant. Consequently, conclusions about the effect of receiver manipulations on ESP performance must be demonstrated by significant interactions between the receiver conditions and the agent - no agent comparison. Very specific statistical analyses are required for such a purpose.

METHODOLOGICAL CONCERNS

Because of the above considerations, the inference that any ESP occurred in the Kreitlers' study will depend upon the degree of certainty we can have that the factors associated with the manipulation of the crucial agent variable - or which covaried with it - did not confound the study with nonpsi variables.

1 On p.207 we learn that when there was to be no agent - the critical comparison condition for all inferences concerning ESP - experimenter B, who was to be with the receiver, learned through the contents of an envelope that he or she was to accompany the subject assigned the role of agent "immediately back to his or her classroom." This procedure was adopted in order that experimenter A (who was normally with the agent) and experimenter C (the coordinator) would not identify "certain trials as perhaps different from others." Unfortunately, this procedure represents a major flaw in the design of the study, for it means that the information available to the receiver's experimenter, as well as the experience of that experimenter just prior to the session, varied systematically for the agent and no agent conditions. What consequences this had for the session we cannot be sure, but this does represent a potentially serious source of confounding for the study. Factors related to this aspect of the design might have resulted in artifactual, nonpsi,

confirmation of the two experimental hypotheses.

2 Potentially important procedural details are not specified in the Kreitlers' report. Who summoned the subject assigned the role of 'agent' on those trials in which there was actually to be no agent? Where was the receiver on the no agent trials while experimenter B (receiver's experimenter) took the 'agent' back to his or her classroom. Was anyone with the receiver during that time? If so, who?

3 Some readers of the Kreitler paper might feel that the study's integrity depends too strongly upon a sense of assurance that the quite elaborate efforts at security were not compromised. While there are certainly some unanswered questions which might be useful to have answered (e.g., details of any efforts, afterwards, to learn whether security was compromised), we prefer to focus our remarks upon aspects of the study which are already clearly problematic based upon information given in the paper.

4 The methodology included variations in the number of receivers tested by the different experimenters and random assignment of sessions (or subjects) to experimenters with the constraint that no experimenter played more than one of the primary experimenter roles (agent experimenter, receiver experimenter, or coordinating experimenter - roles A, B, and C, respectively, in the text of the paper). However, a given receiver experimenter might have tested receivers under only one, under several, or under all of the 12 possible combinations of the agent and receiver variables, and the number tested under each such combination of conditions could vary freely - all this because of random assignment of sessions to experimenters and because the Kreitlers did not control how many sessions each experimenter conducted. It is thus entirely conceivable that factors related to who was the receiver's experimenter (e.g., his/her characteristics) confounded the study through influences upon receiver behavior.

It is unfortunate that the Kreitlers relied heavily upon random methods of trying to control for variables which might have confounded their study. In leading experimental psychology texts such methods are acknowledged to be undesirable whenever it is possible to balance such variables in a direct manner or unless the effects of random assignment can reasonably be assumed to balance matters (as when large numbers of subjects are randomly assigned to each of several experimental conditions). In the study under discussion here, however, any problems associated with the use of random methods (as concerns assignment of experimenters to combinations of the experimental conditions) might have been severely exacerbated by the essentially free variation in the number of sessions conducted by each experimenter.

It would be a little surprising if this random assignment of experimenters to conditions, combined with the free variation in the number of sessions per experimenter, resulted in anything like balanced assignment of experimenters across the combinations of experimental conditions. Since receiver experimenters directly interacted with receivers and certainly influenced the behavior of the latter, any failure to achieve balanced assignment of receiver experimenters across the combinations of experimental conditions might have confounded the study with nonpsi factors.

To avoid potential confounding due to receiver experimenters, each such experimenter should contribute equally to each group involved in a statistical comparison. Random assignment of experimenters to conditions does nothing to insure this, and free variation in sessions per experimenter makes it more difficult - probably impossible - for the condition to be met. To assess possible experimenter confounding the following information is needed which was not supplied in the report: a) the number of sessions (trials) under each of the combinations of experimental conditions which was contributed by each recipient experimenter; b) additional, more specific, breakdowns of this information for the data groupings involved in each of the statistical comparisons from which 'psi' is inferred or from which the effects of recipient conditions upon 'psi' are inferred; and c) statistical analyses appropriate to ascertain whether experimenters were imbalanced relative to the groups involved in any comparison related to 'psi'.

If experimenter imbalance did occur in any of these comparisons, then confounding might have occurred. One might wish to try to salvage the conclusions, nonetheless, by attempting to show that different experimenters did not have differential effects upon receiver behavior, but suitable analysis for that purpose may not be a simple matter and may be impossible (see below). The Kreitlers do report briefly in passing (pp.229-230) that an anlysis (apparently a one-way analysis of variance) was done to assess the effects of recipient experimenter and that its outcome was not significant. They, however, seem to relate that analysis to magnitude of psi effects under different experimenters, and they do not mention the importance of such analyses to ruling out nonpsi interpretations of their data. They do not report the magnitude of the critical statistic, but, in any event, that analysis almost certainly favored a Type II statistical error. As such, it can by no means be seen as ruling out experimenter confounds.

STATISTICAL CONCERNS

1 As was noted above, the only analysis reported by the Kreitlers which has any relevance to possible experimenter effects was (apparently) a one-way analysis of variance which might have favored a Type II statistical error, i.e. a conclusion of no experimenter effect when there was one. The reason the Kreitler's analysis might have favored such an error is that their data showed a very strong main effect of the receiver conditions, and their one-way analysis of variance related to experimenters includes in its error terms (at least part of) the variance from that main effect because effects of receiver conditions have not been partialled out. (The caveat "at least part of" is included because some of that main effect might, itself, be due to experimenter confounding.) Their error term may, thus, be considerably inflated. This could lead to a Type II error. Because of problems related to the freely varying number of sessions contributed by experimenters and the random assignment of experimenters to receiver conditions, it is conceivable that no statistical analysis will prove feasible for disentangling the effects of experimenters from those efects which the Kreitlers really wished to study.

2 The statistical analyses reported in the Kreitlers' paper simply do not provide a basis for a conclusion concerning their second hypothesis, which states that personal-subjective meaning (induced in the receiver) enhances the effect of psi transmission. (At this point the reader should review the second paragraph of the GENERAL PERSPECTIVE section above.) Regardless of which dependent variable is considered as reflecting a possible psi influence, testing of the second hypothesis requires analysis of whether the agent - no agent effect is greater for receivers prepared through personal-subjective meaning than for those who received no special preparation (since preparation through sentence completion may have its own effect on the agent - no agent contrast, an effect which may not be the same as for no preparation, unless that null assumption can be demonstrated

statistically). If it could be demonstrated statistically that the agent - no agent contrast is no different for percipients with no preparation than for those with sentence-completion preparation - using a liberal alpha error such as .10 or .15 for rejecting the null hypothesis (in order to reduce the probability of a Type II error) - then combining the outcomes of those kinds of preparation for the purpose of a contrast with preparation through personal-subjective meaning (as to the degree of agent - no agent contrast) might be considered legitimate. Separate analyses of the agent - no agent effect for receivers getting personal-subjective meaning, sentence-completion, and no preparation would be illuminating, in any case.

The Kreitlers, however, report only significant overall interactions (with several dependent variables) of the agent variable (i.e., agent and no agent), with the levels of treatment of the receiver (i.e., preparation through personal-subjective meaning, through sentence completion, and no special preparation). These analyses are not specific enough to test the second hypothesis since the receiver variable considered in them always has the three familiar levels.

This kind of interaction certainly includes any effect related to the second hypothesis, but it may also include additional effects. Thus, its interpretation, at present, is unclear. For example, suppose that the no-preparation condition resulted in some agent - no agent effect, that the personal-subjective meaning preparation resulted in a larger (though, not necessarily significantly larger) agent - no agent effect as the Kreitlers hypothesize, and that the sentence-completion preparation resulted in no (or a very small) agent - no agent effect as some theorist might suggest because of the syntactic nature of the task. In this hypothetical case the interaction F-value would reflect all these differences, including the difference between the two 'control' or 'comparison' groups, even though this difference is not interpretable as support for the hypothesis that personal-subjective meaning enhances psi effects. A finding that the overall interaction was significant would not imply that any specific component of that interaction was significant. A more detailed analysis, as indicated above, is needed before conclusions can be drawn about the Kreitlers' second hypothesis. (In addition to such an analysis, a table of the 12 cell means would be informative. Despite the various tables and statistical analyses in the Kreitlers' report, the reader is still not provided with information critical to an evaluation of their second hypothesis.)

3 Given the fact that inferences about the occurrence of psi in a study such as this are meaningful only in relation to the agent - no agent contrast, statements such as those on pp.226 and 232 of the Kreitler paper concerned with the (alleged) frequency of manifestation of psi effects (referring to means in their Table 4, e.g., "receivers prepared by personal-subjective meaning manifested psi effects 3 to 4 times more frequently than the other receivers" on p.226) are extremely misleading. The means in Table 4 are simply for the number of instances of a particular type (e.g., direct references to anger) for subjects in specific receiver groups who had an agent. Such agent-condition means, taken alone, cannot possibly sustain the type of inferences which the Kreitlers, as exemplified in the quotation immediately above and often elsewhere in the paper, wish to make from them. Many statements in the paper which lead to conclusions, positive or negative, about psi effects are based upon this same fallacy of trying to assess the effects of a variable upon 'psi' when 'psi' for a given condition is inferred directly (solely) from the single-group mean.

SUMMARY AND CONCLUSION

Because of methodological and statistical weaknesses in the Kreitlers' study of 'Psi transmission of anger and its enhancement through meaning' no conclusions are presently warranted concerning their hypotheses. These weaknesses include, but are not exhausted by, possible nonpsi confounding of the agent - no agent manipulation which is central to inferences concerning psi and a failure to report data and statistical analyses genuinely relevant to testing the hypothesis that psi effects are facilitated by enhancement of personal-subjective meaning in receivers. Additional information is needed and is requested to allow evaluation of the relevance of the outcomes of the Kreitlers' study to the hypotheses supposedly tested. It is our hope that publication of the present critique will facilitate and encourage new, methodologically improved research on the potentially important ideas which the Kreitlers have put before us in their report.

Rex G. Stanford
Psychology Laboratory
SB-15, Marillac Hall
St. John's University
Jamaica, NY 11439
U.S.A.

Ephraim I. Schechter
Psychophysical Research Laboratories
301 College Road, East
Princeton, NJ 08540
U.S.A.

EUR.J.PARAPSYCHOL., V.5, 1984

RESPONSE TO "PSI TRANSMISSION OF ANGER AND ITS ENHANCEMENT THROUGH MEANING? METHODOLOGICAL AND STATISTICAL CONCERNS BY R.G. STANFORD AND E.I. SCHECHTER

Shulamith Kreitler and Hans Kreitler
Tel Aviv University

In their critique of our paper "Psi Transmission of Anger and its Enhancement through Meaning" (Kreitler & Kreitler, 1983) Stanford and Schechter (1984) raise a series of issues that differ in their nature, scope and relevance for ESP research in general. Thus, they could be grouped into classes each of which would include several items that could be discussed together. Yet for the sake of greater clarity and ease in the following our response to each of the claims, we have opted for the serial organization, discussing the claims in the order in which they have been presented by Stanford and Schechter.

1. In their Introductory Section Stanford and Schechter seem to take objection to our style. They apparently hold it against us that we make claims 'boldly' in the title of the paper and elsewhere and "without any indication .... that further research is needed ...". Concerning the 'boldness' it might be advisable to mention that if researchers do not feel confidence in their findings to a degree that suffices for formulating at least a clear title they should not publish their findings. Concerning the need for further research, we would like to refer the reader to pages 231-235 in our paper that deal with some of the most obvious unclarities that require further research.

2. In the first two paragraphs of "General Perspective" Stanford and Schechter suggest that in our study there were no "true ESP scores" and moreover, "in the Kreitlers' study the scores assigned to individual subjects cannot, by the nature of things, be considered as ESP scores. "The nature of things" invoked to separate ESP scores into "True" and "False" is based on what can be called the Rhine tradition of parapsychological research and is at best adequate for test situations that encourage guessing or allow for it. This is not the place for discussing its merits and shortcomings. It suffices to state that this method is anything but adequate for demonstrating impact. If a psychologist, a physicist or a chemist wants to establish whether a particular state or a particular activity of A influences the behaviour of B, he or she carries out experiments that are designed to show whether B reacts differently when exposed to A than when not exposed to A. The significant results of a statistical comparison between the behaviour of a representative sample of Bs exposed to As with the behaviour of Bs in the absence of As are regarded as proof. This generally accepted research methodology served us and others well in psychological and parapsychological experiments. J.B.Rhine accepted it enthusiastically - as he did the study on 'Psi transmission of anger and its enhancement through meaning' - and it did not appear to him to designate the scores obtained by his own method as true scores and our data as devoid of true ESP scores. Now that parapsychology has come of age it may be time for overcoming this harmful paradigm chauvinism reflected in the notion of "true ESP scores" (see also Burdick & Kelly, 1977).

3. The first and major methodological points raised by Stanford and Schechter (see "Methodological Concerns", No.1) stem from highly regrettable typographical mistakes in the published text of the paper. The mistakes consist in mixing up "Experimenter A" and "Experimenter B" in one paragraph. Here is the full text and the corrections in parentheses (Kreitler & Kreitler, 1983, p.207):

> "In trials in which no agent was to participate the envelope for experimenter B (should be: Experimenter A) merely contained the instruction to accompany the subject assigned the role of agent immediately back to his or her classroom. This procedure was adopted in order to prevent experimenters A (should be: B) and C from identifying certain trials as perhaps different from others".

If Stanford and Schechter had read the study a little more carefully than we have read the proofs, they would have immediately grasped the mixup between experimenters A and B, because it is noted in the paper twice (see page 205, 3rd paragraph and 4th paragraph) that "each experimenter participated in the study only in one role (as experimenter A or B or C)" (ibid., p.205). Hence, experimenter A could only be with the subjects who fulfilled the role of agents, and experimenter B could only be with the subjects who fulfilled the role of receivers. Had it been otherwise, the criticism raised by Stanford and Schechter would have been justified. As things are, it remains for us only to express our gratitude to Stanford and Schechter for the opportunity they provided us to correct this typographical error.

4. The second methodological point raised by Stanford and Schechter (see "Methodological concerns", No.2) is based on exactly the same typological error mentioned above (see 3.). Our correction of the error provides the answers to all the questions raised by Stanford and Schechter in this respect, i.e., regardless of whether there was an agent or not, experimenter B was with the receiver on all trials. Thus, there was no difference in procedure either for experimenter B or for the receiver, whether there was an agent or not. Anyway, they did not know about the existence of the agent and there was no way for them to know about it.

5. The third methodological point raised by Stanford and Schechter (see "Methodological concerns", No.3) is in fact no point at all. It is an implied insinuation that "quite elaborate efforts at security" could have been "compromised". We consider such an insinuation out of place in regard to researchers like ourselves at least some of whose previous parapsychological studies have been successfully replicated elsewhere without any intervention on their part. More importantly, in our mind there is no need to believe the integrity of anyone. Science is a public endeavour and experiments are reproduceable events. All our parapsychological studies, including the present one, have been performed with regular subjects not endowed with any particular skills that we know of or that we checked, and under regular conditions that are reproduceable anywhere at any time. Instead of believing or not believing, and instead of insinuating the possibility of "compromise" (a euphemism for fraud?) it would be simpler and certainly more in line with scientific habits to replicate the experiment as faithfully as possible.

6. The fourth methodological point raised (see "Methodological concerns", No.4) concerns the random assignment of experimenters to trials, which according to Stanford and Schechter could have resulted in an unbalanced "assignment of experimenters across the combinations

of experimental conditions", which in turn "might confound the study with nonpsi factors". These concerns derive from the fact that due to space limitations some information relevant for evaluating the findings had to be deleted. We are grateful for the opportunity to provide this information.

Table 1 presents the distribution of experimenters A (the experimenters who stayed with the agents) each separately and all together across the different experimental conditions. Table 2 presents the distribution of experimenters B (the experimenters who stayed with the receivers) each separately and all together across the different experimental conditions. The tables show that both in the case of the five experimenters A and in the case of the five experimenters B the deviations from chance of the distributions of experimenters A and of experimenters B across the experimental conditions - both the receiver conditions and the agent conditions - were not significant. Furthermore, except in the case of one experimenter B (Table 2, experimenter No.5) all other experimenters participated in trials across all experimental conditions. This happened because randomization of trials was strictly adhered to (Kreitler & Kreitler, 1983, page 206, last paragraph). Thus, the data reported in Tables 1 and 2 show that there is no basis for assuming a confounding between particular experimenters and particular experimental conditions.

Moreover, it may be important to repeat what was stated in the original paper, i.e., that no significant differences were found between the five experimenters A and between the five experimenters B in the means of the dependent variables of the study. However, we are glad to add the findings in support of this point which were omitted from the original report because of space limitations. Table 3 represents the receivers' means in the major dependent variables in the trials in which each of the experimenters B (the experimenters who stayed with the receivers) and in those in which each of experimenters A (the experimenters who stayed with the agents) were involved. The one-way analyses of variance show that in no case the differences between the means were significant. Neither were any of the differences between any pair of means significant.

7. The first statistical concern of Stanford and Schechter (see 'Statistical Concerns', No.1) is that the demonstrated psi effects represent a confounding between the effects of the agents and the effects of the experimenters (experimenters A). To our minds, this is

TABLE 1

Distribution of number of trials in which each of experimenters A (who worked with the agents) participated across experimental conditions

Receiver Conditions	Agent Conditions Pers. Meaning	Sent. Comp.	No Prep.	No Agent	Chi-sq
Experimenter A 1					
Pers. Meaning	2	4	6	5	
Sent. Comp.	4	5	4	3	
No Prep.	3	5	5	2	
Total	9	14	15	10	2.23
Experimenter A 2					
Pers. Meaning	3	4	3	2	
Sent. Comp.	6	3	6	4	
No Prep.	3	2	6	2	
Total	12	9	15	8	2.71
Experimenter A 3					
Pers. Meaning	4	2	2	1	
Sent. Comp.	2	1	2	3	
No Prep.	1	1	2	2	
Total	7	4	6	6	.82
Experimenter A 4					
Pers. Meaning	5	1	3	4	
Sent. Comp.	2	3	2	5	
No Prep.	2	4	2	3	
Total	9	8	7	12	1.15
Experimenter A 5					
Pers. Meaning	1	2	2	4	
Sent. Comp.	3	2	3	3	
No Prep.	1	5	1	2	
Total	5	9	6	9	1.76

TABLE 1 (continued)

Total	Agent Conditions					
	Pers. Meaning	Sent.Comp.	No Prep.	Total	No Agent	Total
Experimenters						
A 1	9	14	15	38	10	48
A 2	12	9	15	36	8	44
A 3	7	4	6	17	6	23
A 4	9	7	8	24	12	36
A 5	5	9	6	20	9	29

Chi-square without 'No Agent' =6.44, df=8, n.s.
Chi-square with 'No Agent' =8.42, df=12,n.s.

Total	Receiver Conditions			
	Pers. Meaning	Sent.Comp.	No Prep.	Total
Experimenters				
A 1	17	16	15	48
A 2	12	19	13	44
A 3	9	8	6	23
A 4	13	12	11	36
A 5	9	11	9	29

Chi-square =2.41, df=8, n.s.

TABLE 2
Distribution of number of trials in which each of experimenters B (who worked with the receivers) participated across experimental conditions

Agent	Receiver Conditions			
Conditions	Pers. Meaning	Sent. Comp.	No Prep.	Chi-sq
Experimenter B 1				
Pers. Meaning	5	2	4	
Sent. Comp.	3	3	3	
No Prep.	4	2	4	
No Agent	2	2	5	
Total	14	9	16	2.00
Experimenter B 2				
Pers. Meaning	3	2	4	
Sent. Comp.	2	3	5	
No Prep.	5	1	4	
No Agent	3	4	5	
Total	13	10	18	2.30
Experimenter B 3				
Pers. Meaning	2	6	2	
Sent. Comp.	1	3	4	
No Prep.	2	3	3	
No Agent	3	3	3	
Total	8	15	12	2.11
Experimenter B 4				
Pers. Meaning	6	3	4	
Sent. Comp.	4	2	6	
No Prep.	5	4	3	
No Agent	4	3	5	
Total	19	12	18	1.76
Experimenter B 5				
Pers. Meaning	1	3	1	
Sent. Comp.	2	1	1	
No Prep.	1	1	2	
No Agent	0	2	1	
Total	4	7	5	.87

TABLE 2 (continued)

Total	Receiver conditions		
	Pers. Meaning	Sent. Comp.	No Prep.
Experimenters			
B 1	14	9	16
B 2	13	10	18
B 3	8	15	12
B 4	19	12	18
B 5	4	7	5

Chi square=7.18, df=8, n.s.

Total	Agent Conditions			
	Pers. Meaning	Sent. Comp.	No Prep.	No Agent
Experimenters				
B 1	11	9	10	9
B 2	9	10	10	12
B 3	10	8	8	9
B 4	13	12	12	12
B 5	5	4	4	3

Chi square=1.30, df=16, n.s.

TABLE 3
Comparisons in regard to receivers' means between experimenters who worked with the receivers and between experimenters who worked with the agents

Variable	Receivers' Means *					F	Sig. Group Diffe- rences **
Direct references to	1.80	2.10	1.93	1.42	1.50	1.01	None
anger in reporting	1.78	1.86	1.70	1.68	1.70	.86	None
Indirect references	1.82	1.78	1.81	1.60	1.80	.74	None
to anger in reporting	1.84	1.68	1.64	1.62	1.92	.61	None
Total of references	3.62	3.88	3.74	3.02	3.30	1.12	None
to anger in reporting	3.62	3.54	3.34	3.30	3.62	.96	None
References to anger	2.87	3.00	3.07	2.76	2.40	.44	None
in sentence comple- tions	2.95	2.72	3.09	2.83	2.88	1.42	None
Difference in eva-	-1.41	-1.43	-1.49	-1.42	-1.45	.53	None
luation of anger before and after reporting	-1.48	-1.42	-1.39	-1.40	-1.45	.63	None

*) For each variable the first row reports the data when each of the five experimenters who worked with the receivers (experimenters B) was involved and the second row reports the data when each of the five experimenters who worked with the agents (experimenters A) was involved. Experimenters A were not the same individuals as experimenters B. The number of subjects with whom experimenters B worked were 39, 41, 35, 49 and 16, respectively; the number of subjects with whom experimenters A worked were 38, 36, 17, 24 and 20 respectively.
**) Comparisons of pairs of means were done both by the Scheffé method and Duncan's range test.

first and foremost not a statistical concern but a theoretical concern that depends partly on the experimental design. Since the agent always stayed with an experimenter and since the experimenter was present during the role playing of the psychodramatic scenes designed to evoke anger, it is possible that the demonstrated psi effect is due either only to the agent or only to the experimenter who was with the agent or to some interaction of both agent and experimenter whose nature we do not know. The possibility that the psi effect is perhaps due only to the experimenter is enhanced by the following findings of the original study: the psi effect was not found to be related to the agent's seriousness of attitude toward the enactment of anger or toward feeling and transmitting anger or to the degree to which the agent made overt movements denoting anger or seemed to be immersed in the psychodramatic enactment or seemed actually to feel anger during the psychodramatic scene or admitted to having felt anger (Kreitler & Kreitler, 1983, p.229). All these findings suggest that although the experimenter was busy during the session and was not enacting anger the mere observing of the agent's enacted anger or in some form thinking about anger or the anger-provoking scene could have sufficed for the psi effect. Thus, it is conceivable that the experimenter could have been the actual transmitter of the anger to the receiver. There is no way in which we could find out who in this experiment is responsible for the effect - the agent or the experimenter A - not merely because agent and experimenter A always appear together (when the agent does not enact anger experimenter A also is not exposed to it) but mainly because we do not know how the transmission takes place. Thus, in whatever way we subdivide the data we will always end up with the unit agent-experimenter A and will be unable to establish who is the source of the observed impact.

However, paradoxically, these remarks do not indicate that there is a confounding between the effect of the experimental manipulation and the effect of the experimenter. Strictly speaking, confounding occurs when two or more different effects are so entangled that there is no way to specify which effect is responsible for the observed impact. Yet, in our case the effect is the same, i.e., psi transmission of anger. The unclairity refers merely to the identity of the agent: Is it the subject to whom the role of the experimenter was assigned or is it the agent-experimenter pair? Likewise, it may be observed that a similar problem exists in regard to disentangling telepathy from claivoyance.

Nevertheless, it is possible to study separatly the effects of actual agents and of experimenters. At present this problem has to be added to the list of problems that await research in the future.

8. The second statistical concern (see 'Statistical concerns', No.2) refers to the testing of the second hypothesis about the effect of preparing the receivers by personal-subjective meaning. Stanford and Schechter claim that the analyses reported in our paper were not sufficiently specific for testing the hypothesis. Our analyses were based on comparing receivers who got preparation by personal-subjective meaning with those who got no preparation or preparation by sentence completion. Stanford and Schechter suggest that only analyses based on comparing agent-no agent data for each receiver condition separately could provide information relevant for the confirmation or rejection of the hypothesis. Such comparisons are presented in Table 4 for each of the eight dependent variables involved in assessing the impact of psi. Each comparison is based on testing the following null hypothesis

$$(\mu_1 + \mu_2 + \mu_3)\ /\ 3 = \mu_4$$

whereby μ_1, μ_2 and μ_3 represent the means of the variable under a given receiver condition when the agent is present (μ_1 -when the agent gets preparation by personal-subjective meaning, μ_2 -when the agent gets preparation by sentence completion, μ_3 -when the agent gets no preparation) while μ_4 represents the mean of the variable under a given receiver condition when there is no agent.

Table 4 shows that in each case the only comparison that yields significant results is the comparison which refers to the condition when the receiver got preparation by personal-subjective meaning. Hence, these analyses support fully the second hypothesis of the study.

In addition, Table 4 presents for each of the eight variables the 12 means which Stanford and Schechter found lacking in the original paper.

9. In the third and last of their statistical concerns (see 'Statistical concerns', No.3) Stanford and Schechter claim that many of the statements we made about psi were solely on the basis of single-group means. They actually give only two examples for this, both of which belie the very claim. Thus, the statements on p.226 and p.232 (original report) do not refer to absolute frequencies but to

TABLE 4
Mean comparisons of agent vs no agent conditions
in each receiver condition separately

Variable	Receiver conditions	Agent Conditions (a) Pers. meaning	(b) Sent. comp	(c) No prep.	(d) No agent	(e) Mean of (a)(b)(c)	F (e)&(d)
Direct ref to anger in reporting	Pers.Mean.	3.73	3.33	3.07	.46	3.38	7.53**
	Sent.Comp.	1.13	.80	.60	.40	.84	.16
	No Prep.	1.40	1.20	.60	.20	1.07	.63
Indir.ref to anger in reporting	Pers.Mean.	3.87	3.80	3.20	.13	3.62	9.48**
	Sent.Comp.	.73	.60	.93	.47	.75	.06
	No Prep.	.67	1.00	1.00	.33	.89	.25
Direct & indir.ref to anger in reporting	Pers.Mean.	7.60	7.13	6.27	.53	7.00	12.34**
	Sent.Comp.	1.86	1.40	1.53	.86	1.60	.16
	No Prep.	2.07	2.26	1.60	.53	1.97	.61
Refer.to anger in sent.comp	Pers.Mean.	3.86	4.00	3.80	.86	3.89	7.58**
	Sent.Comp	2.40	2.26	1.87	.40	2.18	2.59
	No Prep.	2.73	2.53	2.40	.60	2.55	3.14
Evaluat. of anger after reporting	Pers.Mean.	2.73	2.73	2.73	1.40	2.73	7.34
	Sent.Comp.	2.40	2.13	2.00	1.53	2.18	1.75
	No Prep.	2.14	2.06	1.73	1.26	1.98	2.09
Differen-rence in eval. of anger before&after reporting	Pers.Mean.	-1.68	-1.51	-1.66	- .28	-1.62	5.91*
	Sent.Comp.	-1.13	-1.05	- .81	- .46	-1.00	.96
	No Prep.	-1.06	- .89	- .66	- .08	- .87	2.05

* p<.05
** p<.01

comparisons which necessarily involve more than one group mean. More importantly, the differences between the means referred to - in these and other cases - have been examined for significance. Thus, when we write that the "receivers prepared by personal-subjective meaning manifested psi effects 3 to 4 times more frequently than the other receivers" (p.226) this statement is based on the significant mean differences (please see original paper, Table 4, last column). Actually the range of the differences in the three major dependent variables mentioned in this context is 3.16-4.38. The comparisons are based on means of receiver groups when the agent is present because these are the conditions relevant to the question we posed. If one wants to compare the frequencies of psi effects, it makes no sense to compare the frequency when no agent is present. The latter is what Stanford and Schechter seem to have in mind. It would have provided much more dramatic estimates of the effect (e.g., 8.11 in the case of direct references to anger, 29.77 in the case of indirect references to anger) but these do not seem to be relevant to our purpose in that context. Incidentally, it may not be superfluous to repeat at this point what was mentioned also in the original report, i.e., that whenever group mean comparisons were made we applied the Scheffé method, which is an a posteriori method (while an a priori less strict method would have often been justified) and moreover is the most demanding of the a posteriori methods. To quote Winer (1971, p.201): "The Scheffé method is clearly the most conservative with respect to a type I error; this method will lead to the smallest number of significant differences. In making tests on differences between all possible pairs of means it will yield too few significant results". These remarks make it abundantly clear that we did not tend to see a difference where actually there is none.

10. In their Summary and Conclusions Stanford and Schechter write: "Additional information is needed and is requested to allow evaluation of the relevance of the outcomes of the Kreitlers' study to the hypotheses supposedly tested". Owing to poor formulation the end of this sentence could have two readings. It could mean that they consider the hypotheses as not confirmed because they were poorly tested, in which case it should have been "supposedly confirmed", or - which is more in line whith the actual formulation - it could include the insinuation that we in fact did not carry out the study. For the benefit of the reviewers we assume that it is the first reading that corresponds to their intention. If so, we hope that we have shown by our responses to each of the raised points that the hypotheses were not only "supposedly tested" but were also actually

confirmed.

ACKNOWLEDGEMENTS

Our thanks are due to Mr. Joseph Glickson of the Department of Psychology, Tel Aviv University, for his insightful suggestions.

REFERENCES

Burdick, D.S., Kelly, E.F. 'Statistical methods in parapsychological research. In: Wolman B. (Ed.), 'Handbook of parapsychology', New York: Van Nostrand Reinhold, 1977.

Kreitler, S., Kreitler, H. 'Psi transmission of anger and its enhancement through meaning', E.J.P., 1982, 4, 199-241.

Stanford, R.G., Schechter, E.I. 'Psi transmission of anger and its enhancement through meaning? Methodological and statistical concerns', E.J.P., 1984, 5,

Winer, B. 'Statistical principles in experimental design', 2nd edition, New York: McGraw-Hill, 1971.

H. Kreitler & S. Kreitler
Tel Aviv University
P.O. Box 39040
Tel Aviv 69978
Israel

PA CONVENTION 1984

The 27th Annual International Convention of the Parapsychological Association will be held from Monday August 6th, 1984, to Friday, August 10th, 1984, in Dallas, Texas, U.S.A., at the Southern Methodist University. Activities start on Sunday evening with a 'Welcome to Texas' party.

The Program Committee is chaired by G.F. Solfvin.

For further information, contact Martha Dealey, 6927 Leameadow, Dallas, Texas 75248.

EUR.J.PARAPSYCHOL., V.5, 1984

AUDIATUR ET ALTERA PARS
A TRIBUTE TO A CLOSE FRIEND: PIET HEIN HOEBENS (1948-1984)

Gerd H. Hövelmann

"Life's too long
(as the Lemming said)..."
(Jethro Tull, 1978)

I cannot remember having ever received a news that meant such a deep and lasting shock to me as that of the untimely death of Piet Hein Hoebens on October 22, 1984, shortly after his 36th birthday. First reactions I have received to date (November 5, 1984) from other collaborators and friends of Piet Hein, such as Prof. Marcello Truzzi and Dipl.-Psych. Eberhard Bauer, convey very similar feelings. Nevertheless, Piet Hein's death did not meet us completely unprepared, because at least Marcello Truzzi and I knew that he had felt exceedingly unwell for quite some time and that he expected to die before long. In his last letter to me, mailed about a week before his death, Piet Hein wrote:

> "I have reasons to suspect that my 36th birthday may have been my last birthday on this side of the Grave. I am not sure of it, but it is a distinct possibility ... Of course, the most desirable thing would be that I simply continue living, but some contingencies are quite beyond human control."

Piet Hein was a card-carrying skeptic in matters parapsychological and a prominent and extremely influential associate (and Dutch representative) of the American-based skeptical Committee for the Scientific Investigation of Claims of the Paranormal (CSICOP). Despite these facts, I believe that his death means an even more serious loss to responsible parapsychology (about which I will say more presently) than it does to the community of the critics of the field.

Piet Hein was a man of broad learning, and he had many quite different talents. He was born on September 29, 1948, in Utrecht, The Netherlands. He was a connoisseur of and had a life-long interest in art and classical music which, as he once recalled, was the reason for considerable trouble he had with his "uncouth fellow pupils" when he was an eight-year-old school boy. After finishing school, Piet Hein began to study law, which he gave up after a few semesters, however, when he was offered the opportunity to join the staff of one of the leading Dutch newspapers, 'De Telegraaf', in Amsterdam. He remained on that staff for 13 years until his death, and he became a prominent and well-reputed investigative journalist and editorial writer. As the editors of 'De Telegraaf' noted in their obituary ("Piet Hein Hoebens overleden", De Telegraaf, October 25, 1984, p.T 3) and as anyone who has read his numerous publications in professional journals will confirm, he was an exceptionally talented writer ("begenadigd met een groot schrijverstalent"). Piet Hein was quite fluent in several languages: French, German, and, of course, Dutch; his English was almost indistinguishable from that of a native speaker. In addition, he was able to read more than half a dozen further languages, including Danish, Swedish, Spanish, Italian as well as classical Greek and Latin. Also, he was a gifted cartoonist, as can be seen from the cartoons which were published occasionally in the 'Zetetic Scholar' and in the 'Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie'. However, the vast majority of his funny cartoons have never been published. Piet Hein never lost his quiet, gentle sense of humor, not even in that rather lugubrious last letter he wrote to me.

I first got in touch with Piet Hein in March of 1981 shortly after he had conclusively demonstrated in the pages of the 'Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie' the untrustworthiness of the scientific work of the late Prof. Willem Tenhaeff. (3) (note 1). We soon realized that there were a number of common features in our respective views on the scientific and epistemological status of parapsychology and on the relations between parapsychologists and their critics. This marked the beginning of very extended and detailed discussions (both personally and by correspondence) on these questions in the course of which we not only became colleagues and collaborators but close personal friends. Our extensive correspondence of the last three and a half years amounts to several hundred pages. And we ended up working on several joint projects. Except for some of my academic teachers, most notably Marburg philosophers Prof. Peter Janich and Dr. Holm Tetens, there was hardly anyone who had greater impact on my own views of science and philosophy in these past few years. And he was not even a scientist; nor was he a philosopher! Nevertheless, he was well-acquainted with the philosophical literature (from David Hume to Imre Lakatos), and his science-philosophical insights frequently were more lucid than much of what I found in the writings of professional philosophers and in the pages of specialized philosophical literature. Piet Hein had intended to explain his philosophical views in a book on parapsychology he was working on and which was to contain a long chapter on 'Parapsychology and the Philosophy of Science'. Although we also agreed in our rejection of racism, anti-Semitism, and other questionable cultural value systems, there always remained a profound disagreement between us as far as the question of general political orientations is concerned. In these matters, he preferred a position that was more conservative than the one I feel able to subscribe to.

Piet Hein, who once wrote that he "generally much prefer(red) Truzzi's approach to that of the CSICOP," was a man of impeccable intellectual integrity and self-discipline, virtues which also came to the fore in his involvement in controversies around parapsychology. Parapsychologists have frequently complained (and not completely without justification) that some of their critics are not sufficiently familiar with the work they are attacking, that they are more interested in debunking and discrediting than in scientific truth, that they portray dedicated experimentalists as credulous occultists and tend to base their criticisms on caricatures of parapsychological research. However, Piet Hein was a rara avis in these controversies around parapsychology in that he (like a few others, most notably Prof. Marcello Truzzi and Prof. Ray Hyman) did not at all fit this stereotype of the uncritical critic which some parapsychologists seem to have taken quite a fancy to. On the other hand, it did not escape Piet Hein's attention that leading parapsychologists are well aware of the difference between responsible and irresponsible types of skepticism. In a recently established Belgian parapsychological journal, 'Psi-Forum', he wrote:

> "Sensible parapsychologists (like Martin Johnson, John Beloff, Brian Millar, Eberhard Bauer, Douglas Stokes, Charles Akers, Sybo

> Schouten, Gerd Hövelmann, Susan Blackmore, Robert Morris, Walter von Lucadou, and the SRU-quintet in Eindhoven) will not be misled by this sort of stereotypes." (24: p.17) (note 2).

Many parapsychologists have realized and acknowledged the fact that Piet Hein played a very special role in the continuing parapsychology and its legitimacy as a scientific endeavor. Dr. John Palmer, for instance, called him "my Dutch colleague" (Palmer, 1983, p.39, footnote), thus indicating that the Hoebens-type skepticism is not even remotely comparable to that which Palmer was going to criticize in that article. In 'Fate', a magazine that certainly none will seriously suspect of being excessively sympathetic to CSICOPers, Jerome Clark wrote:

> "Hoebens is a first-rate investigative journalist who uses facts and documentation, not propaganda and speculation, to make his case. Hoebens, whose writings are appearing with increasing frequency in this country, is as appealing a skeptical commentator on the paranormal as anyone currently on the scene - a man of such uncommon sense, integrety and good humor that one wishes there were more like him on both sides of the paranormal debate" (Clark, 1983, pp. 93-94).

Reflecting on his reputation among responsible parapsychologists, Piet Hein himself wrote in an undated (January 1984) letter to selected "Dear parapsychologists":

> "Although I am a skeptic, a non-believer in ESP, PK and Poltergeists, a friend of Dr. E.J. Dingwall, a defender of Project Alpha, an aficionado of Hume's Essay on Miracles and an unrepentant member of the Committee for Scientific Investigation of Claims of the Paranormal I am told that many parapsychologists think of me as a fair, reasonable and open-minded critic. Nothing is as flattering as being held in esteem by one's opponents, so naturally I am proud of my reputation".

Despite his professed skepticism, Piet Hein frequently and publicly (e.g., 3, 12, 15, 23, 24, 32, 33) defended the legitimacy of the parapsychological research programme (he even served as a subject in several parapsychological experiments; for instance, cf. (1)), and he made no secret of his sometimes profound disagreements with certain of his fellow-skeptics. Many of his (mostly private) criticisms of fellow-CSICOPers and other critics of parapsychology are often sharper and sometimes more to the point than anything parapsychologists have written about what they perceived as unfairness and uninformedness on the part of the critics. His many semi-public memos were instrumental in bringing the CSICOP/Mars Effect Controversy back again to the level of rational argumentation, and his last article on this affair (19) drew praise from Marcello Truzzi and Michel Gauquelin as well as from Ken Frazier and Paul Kurtz, which must have been the first time in the history of that affair that these people agreed about anything (note 3).

Even more important than Piet Hein's public activities were his numerous and many-faceted behind-the-screen efforts to secure fair and rational treatment of parapsychology (as he once remarked in a personal conversation, "Skeptics who lump Beloff and Berlitz together are deliberately deceiving the public"). Even now that he has passed away, I am not yet able to reveal some of these activities. He made many (and many successful) attempts to install improved and reasonable communications between parapsychologists and critics; he organized international support (among skeptics!) for the Parapsychology Laboratory at Utrecht State University and for the Lehrstuhl für Psychologie und Grenzgebiete der Psychologie at Freiburg University which both were in danger of being hit by budget cuts; the vast majority of fair or even favorable statements about parapsychology which appeared in recent volumes of the 'Skeptical Inquirer' either were authored by himself or can be traced back to his direct or indirect influence. He initiated and chaired a meeting of German parapsychologists (Bauer, von Lucadou, Hövelmann) with a leading German skeptic, Prof. Irmgard Oepen, which was held in my appartment in Marburg in November of 1982. The catalogue of nine basic statements to which all of us felt able to agree and which was subsequently published in the 'Skeptical Inquirer' (Frazier, 1983), soon acquired the sobriquet "Marburg Manifesto", and it was extensively used by Dr. Stanley Krippner for his excellent presentation before the 1984 Annual Meeting of the American Association for the Advancement of Science (Krippner, 1984). Others of his activities in support of reasonable and responsible parapsychology will become more evident as time goes by. Piet Hein clearly realized that he had a special responsibility to do whatever he could to prevent his extremely valuable files, including many important documents and his extensive professional correspondence, from getting lost, literally or figuratively. Suffice is to say here that he made appropriate arrangements before his death.

In recent month, the sincerity of Piet Hein's liberalism in matters parapsychological has been questioned by several parapsychologists who, however, did not take the trouble to contact him directly and to give him a chance to respond to their allegations. These charges mostly concerned his debunking work on some German cases and experiments (for instance, c.f. 16, 18, 21), including his devastating critique of Bender's famous Pirmasens Chair Test with Gerard Croiset (31). I may note here the record that almost all of Piet Hein's work on these German investigations and the way he handled and examined the relevant material has been closely observed by both Eberhard Bauer and myself. As far as I am concerned - I am not entitled to speak for Eberhard Bauer (though I am sure that he will agree with me) - I was unable to detect even the slightest sign of unfairness, overinterpretation, or misinterpretation in this work. As for other criticisms and charges of unfairness, all those which so far have been brought to my attention can easily be refuted on the basis of Piet Hein's professional correspondence.

Piet Hein did not attach importance to formal positions; he was quite satisfied with being able to influence the course of events in an unofficial and non-public capacity. Thus, he did not accept invitations from both the 'Skeptical Inquirer' and the 'Zetetic Scholar' to become Consulting Editor of the former and Associate Editor of the latter. Moreover he remained quite unimpressed when one or the other of his opponents tried to substitute appeals to authority and academic degrees and credentials for sound argumentation. As he once put it pertinently in a letter to me:

> "Being more interested in arguments than in formal qualifications ..., I think that to be entitled to a degree means the same thing as being able to prove that you are entitled to a degree. That is what degrees are for."

From somewhat different starting points, Piet Hein and I approached what we were told is a demarcating line dividing parapsychologists from their critics, but what we found was that, to all intents and purposes, this demarcating line is imagery, non-existent! Responsible parapsychologists and responsible critics are doing essentially the same work, and in the last resort they are pursuing the same goals, improving science, that is. To describe Piet Hein's views on the nature of controversies around parapsychology, I can do no better than use his own words. In what is probably his last article (on 'Fraud and Selfdeception in Parapsychology'), he wrote (33) (note 4):

> "As a habitual skeptic, I suspect 'paranormal phenomena', as conventionally understood, of non-existence. However, I am clearly aware of the dangers of sceptical dogmatism and prefer to see the debate over 'psi' as an intellectual game the outcome of which cannot yet be predicted with any degree of confidence. The sceptical position is based on certain assumptions that may eventually become obsolete. I further concede the point that several well-known critics are insufficiently familiar with the subject and seriously underestimate the case for parapsychology. For the purpose of the present essay I will resign myself to the traditional, stereotypical proponent/sceptic dichotomy which is in fact a gross simplification of the actual state of affairs ...
>
> Several influential critics of parapsychology alas delight in misrepresenting the field. According to their polemical writings, parapsychology is medieval superstition parading as modern science. Its ultimate objective is the overthrough of the reign of Reason and a restoration of the dark ages of magical belief. In order to conceal their irrationalist aims, the parapsychologists have concocted transparently phoney 'proofs' which of course will evaporate instantly as soon as a cardcarrying sceptic takes the trouble to subject them to a cursory examination. The parapsychologists are unable to see the obvious, as they are blinded by the metaphysical prejudice."

Readers who suspect that here Piet Hein might have exaggerated skeptical misconceptions and prejudice should very careful look at parts of the recent book by Dr. Jame Alcock (1981) and at Prof. Mario Bunge's most recent criticism of parapsychology (Bunge, 1984, esp. pp. 42-44); Bunge obviously does not have more than a very remote idea of the field he is writing about, and what is worse, he apparently does not even want to become at least moderately familiar with the field and the work he condemns as 'pseudoscientific'. Piet Hein continues:

> "Although several individual parapsychologists convey the impression of being obsessed by the desire to conform as closely as possible to the above stereotype it would be grossly unfair to claim that this is true for the parapsychological community as a whole.
>
> The fundamental error in the propaganda of the extreme sceptics (apart from their tendency, seriously to underestimate the scientific case for 'psi') is their tacit assumption that a

> committment to parapsychology implies a strong belief in supernatural forces and that, therefore, only outsider sceptics are capable of a critical assessment of claims of the paranomal. In fact, the constitutions of the major parapsychological organizations have traditionally been non-committal as to the authenticity of the paranormal phenomena. While most parapsychologists accept 'psi' as real or very probably real, others have dissented without their dissent leading to excommunication ...
>
> The reality of the psi debate is far more complex than is apparent from the polemical writings pro and con, with their monotonous emphasis on the stereotypical believer/sceptic dichotomy. Three of the sharpest critics of parapsychology alive today, Charles Akers, Susan Blackmore and Gerd Hövelmann, are themselves members of the Parapsychological Association and, ipso facto, recipients of ultra-sceptical abuse ...
>
> Rather than indulging in the futile passtime of proclaiming the superiority of one's own insight, believers and unbelievers should view their expectations as stakes in a game-like, Lakatosian duel of 'research programmes'. Here, the proponent predicts that increasingly sophisticated research will result in a progressive accumulation of findings supportive of the psi hypothesis, whereas the sceptic predicts a progressive erosion of the parapsychological evidence as it becomes increasingly amenable to 'naturalistic' explanations."

The most important legacy, then, Piet Hein left to responsible parapsychologists and cooperation between both 'groups' (and there are reasons to doubt that there really are such things as parapsychological and skeptical camps), based on mutual respect, can give us any hope for a scientific solution to the problems at issue. This implies that parapsychologists are candid about the weak points and the speculative nature of many of their theoretical constructions and about possible loopholes in the design of their experiments, and that critics take the trouble to make themselves sufficiently familiar with the work they are going to subject to skeptical scrutiny. Appeals to 'higher' insights and unfounded claims of superiority on either side of the psi debate, do not take us anywhere. Indeed, responsible parapsychology has nothing to lose and much to potentially gain by rational cooperation with its nominal opponents. The same applies, mutatis mutandis, to responsible critics. That, then, is the task we are left with. However, even if we achieve these goals, I fear, responsible skepticism will never be the same without Piet Hein; nor will rational parapsychology; nor will I.

Piet Hein will be missed by many who are involved in the debates around parapsychology. Those of us who, like Marcello Truzzi, Eberhard Bauer, and myself, were not only collaborators, but also close personal friends of Piet Hein, are not the only ones to whom his death is a grave loss. The most genuine tribute any of us can offer Piet Hein is to find a way in our work to realize the ideals he stood for: Trustworthiness, meticulousness, impeccable intellectual integrity, and an obsession with doing his various jobs as good as they possibly could be done. Those of us who knew him well, and those whom he counted among his friends, were priviledged. We are the poorer for his departure. Piet Hein is survived by his young wife, Liesbeth, to whom we extend our sympathy.

NOTES

1. Numbers in paraenthesis refer to the Select Bibliography of P.H. Hoebens' writings, below.
2. Translation is mine. The SRU-quintet in Eindhoven, which Piet Hein refers to in this quotation, consists of Heyme Breederveld, Jeff C. Jacobs, J.A.G. Michels, R. Pare , and A.C.M. Verbaak.
3. As Prof. Truzzi tells me, the forthcoming issue of the 'Zetetic Scholar' will contain Piet Hein's detailed review of the latest book by Michel Gauquelin.
4. Another article by Piet Hein on 'Psychic detectives' written from a zetetic perspective, was in its almost final form at the time of his death. Piet Hein made arrangements with Prof. Truzzi, and he instructed him to rearrange the material and complete that paper. It will appear in Kurtz, P. (ed.), 'A Skeptics Handbook of Parapsychology', Buffalo, N.Y.: Prometheus Books. So the article which I have quoted from in the text is not the last one he wrote.

REFERENCES

Alcock, J. 'Parapsychology: Science or magic?', Oxford, Pergamon Press, 1981.

Bunge, M. 'What is pseudoscience?', Skeptical Inquirer, 1984, 9, 36-46.

Clark, J. 'Parapsychologists, critics agree to consensus statement', Skeptical Inquirer, 1983, 7, (4), 4-6.

Jehtro Tull, 'Moth', on the record 'Heavy Horses', Los Angeles, Chrisalis Records, 1978.

Krippner, S. 'Parapsychological phenomena: A topic to investigate', Paper presented at the 150th Annual Meeting of the American Association for the Advancement of Science, New York, May 29, 1984 (36 pages).

Palmer, J. 'In defence of parapsychology: A reply to James E. Alcock', Zetetic Scholar, 1983, 11, 39-70.

A SELECT BIBLIOGRAPHY OF PIET HEIN HOEBENS'S WRITINGS ON PARAPSYCHOLOGY AND RELATED MATTERS

1. 'How I was debunked', Skeptical Inquirer, 1979-1980, 4, (2), 64-66.

2. 'Comments (on a paper by Edward W. Karnes et al.)', Zetetic Scholar, 1980, 7, 131-132.

3. 'Vom Lob der Genauigkeit in der Parapsychologie' (translated by Eberhard Bauer), Zeitschrift f. Parapsychologie u. Grenzgebiete d. Psychologie, 1980, 22, 225-234.

4. 'Gerard Croiset: Investigation of the Mozart of "psychic sleuth"-Part I', Skeptical Inquirer, 1981, 6, (1), 17-28.

5. 'Gerard Croiset, le medium qui n'a jamais faili' (translated by Claudine Braine), Les Cahiers Rationalistes, 1981, No. 369, 245-250.

6. '"Ik was iemand anders"', De Telegraaf, 1981 (september 21), T 29.

7. 'The mystery men from Holland. I: Peter Hurkos' Dutch cases', Zetetic Scholar, 1981, 8, 11-17.

8. 'Prof. Chari licht kritiek op boek over reincarnatie toe', de Telegraaf, 1981 (December 19).

9. 'Croiset and Professor Tenhaeff: Discrepancies in claims of clairvoyance', Skeptical Inquirer, 1981-1982, 6, (2), 32-40.

10. 'Mystery men from Holland. II: The strange case of Gerard Croiset', Zetetic Scholar, 1982, 9, 20-32.

11. 'Comments (on a paper by Patrick Curry)', Zetetic Scholar, 1982, 9, 70-71.

12. 'Die Legitimität des Unglaubens' (translated by Eberhard Bauer), Zeitschrift f. Parapsychologie und Grenzgebiete d. Psychologie, 1982, 24, 61-73 (reprinted in Bauer, E. and von Lucadou, W. (eds) 'Spektrum der Parapsychologie', Hans Bender zum 75. Geburtstag, Freiburg i.Br., Aurum Verlag, 1983, 118-130).

13. 'The mystery men from Holland. III: The man whose passport says clairvoyant', Zetetic Scholar, 1982, 10, 7-16.

14. 'Comments (on a paper by Ron Westrum)', Zetetic Scholar, 1982, 10, 124-126.

15. 'Cambridge Centenary of psychical research: Critics heard, encouraged to cooperate', Skeptical Inquirer, 1982-1983, 7, (2), 2-4.

16. 'Kritische analyse van Bender's 'Kriegsprophezierungen' (sic)', Prana, Tijdschrift voor geestelijke verruiming en randgebieden der Wetenschappen, 1983, No. 33, 94-98 (the editor made some distorting and non-authorized changes).

17. 'Sprekende dieren houden geleerden aan de praat ...', De Telegraaf, 1983 (August 6), T 25.

18. 'Parapsychologie kontrovers: Die Erfasser des Unfassbaren' (translated by Rüdiger Runge), Psychologie heute, 1983, 10, (9), 64-67.

19. 'Some further reflections on the Mars Effect Affair', Zetetic Scholar, 1983, 11, 25-28.

20. 'Comments (on a paper by Gerd H. Hövelmann)', Zetetic Scholar, 1983, 11, 145-147.

21. 'Vernebelungstaktik' (translated by Rüdinger Runge), Psychologie heute, 1983, 10, (12), 19.

22. 'The Hume game', in Roll, W.G., Beloff, J. & White, R.A. (eds), Research in Parapsychology 1982, Metuchen, N.J. and London, Scarecrow Press 1983, 15-17.

23. 'Sense and nonsense in parapsychology: Parapsychology spans a spectrum from sophisticated skepticism to pro-psi gullibility', Skeptical Inquirer, 1983-1984, 8, 121-132.

24. 'Skeptici bijeen', Psi-forum, 1984, 1, (1), 17-20.

25. 'Coming of age in Buffalo: Reflections on the 1983 CSICOP Conference', Skeptical Inquirer, 1984, 8, 202-206.

26. 'Netherlands: Psychic surpluses and a call for international cooperation', Skeptical Inquirer, 1984, 8, 227-229.

27. 'De balk in eigen oog', SRU Bulletin, 1984, 9, 5-12.

28. 'Skeptische concilie te Buffalo', Psi-Forum, 1984, 1, (2).

29. 'Een antwoord aan Dr. D.J. Bierman', SRU Bulletin, 1984, 9, 72-75.

30. 'Letter to the Editor (Comment on a statement by Jürgen Keil)', Zetetic Scholar, 1984, 12, in press.

31. 'Abschied von Pirmasens. Eine kritische Nachuntersuchung eines erfolgreichen ASW-Experiments', Zeitschrift f. Parapsychologie und Grenzgebiete der Psychologie, 1984, 26, in press.

32. 'Grenzgebiete der Medizin und die Verantwortung der Parapsychologen' (translated by Gerd H. Hövelmann), in Oepen, I. (ed) 'Aussenseitermethoden in der Medizin', Darmstadt, Wissenschaftliche Buchgeselschaft, 1985, in press.

33. 'Fraud and self-deception in parapsychology', in Duerr, H.P. (ed), 'Authenticity in science', English and German editions, in press.

Gerd H. Hövelmann,
Rollwiesenweg 42,
3550 Marburg/Lahn,
West Germany.

EUR.J.PARAPSYCHOL., V.5, 1984

IMPROVEMENT OF THE 'PROBABILISTIC PREDICTOR PROGRAM' OF TART AND DRONEK FOR TESTING RANDOM TARGET GENERATORS

Z. Vassy

THE 'PROBABILISTIC PREDICTOR PROGRAM' OF TART AND DRONEK

In most of ESP experiments sequences of random numbers are used to determine the order of targets. The randomicity or possible non-randomicity of the target sequence is crucial especially in experiments where the percipient receives immediate feedback about the correct target after every trial because some regularities in the sequence of targets may lead to spurious hits by the use of (largely unconscious) inference strategies. Tart and Dronek (1982) give a clear and convincing analysis of this possibility and point out that the usual chi-square test for deciding about general randomness is a rather weak tool if one wants to estimate a measure of predictability of a given sequence or type of sequences. They present an algorithm, the 'Probabilistic Predictor Program', that simulates a subject making stochastic inference about target order, and they characterize the degree of non-randomicity of the random target generator by the performance (score) of the predictor program.

The Tart-Dronek (TD) algorithm works in the following way. There are counters that store the number of previous occurrences of all individual targets, target doublets, triplets, etc. up to sextuplets (in their particular implementation; other choices are, of course, possible as well). For example, if the targets are the digits 0,1,...,9 and the triplet (2,7,5) occurred six times before the Nth trial, then the counter C(2,5,7) will contain the number six at this stage. Now assume that the Nth target was 7 and the (N-1)th one was 2. Considering triplets only, which number is the most likely to follow? The answer is found by looking through the counters C(2,7,X) for X=0,1,...,9 and choosing the particular X_o for which the content of the counter is the largest: $M(N,\text{triplets})=\max_{(X)} C(2,7,X)=C(2,7,X_o)$ and PREDICTION(N,triplets)=X_o. This is quite straightforward and obvious. But what can we do if for singlet, doublet, triplet etc. subsequences the candidate targets are not the same? For example, it may turn out that up to the Nth target the most frequent singlet target was 4, that seven was followed most frequently by 9 etc.. How can we decide which apparent bias is the strongest? The TD choice is to compare the statistical significances of deviations of M(N,k-tuplets) from their expected unbiased values. The apriori (unbiased) probability of any particular k-tuplet in one trial is $1/T^k$ where T is the number of targets (in our example T=10). In N trials the probability of that any particular k-tuplet occurs exactly j times is

$$\binom{N}{j}(1/T^k)(1-1/T^k)^{N-j}$$

by the binomial distribution. Hence the probability that the most frequent k-tuplet occurs M(N, k-tuplets) times or more, i.e. the statistical significance level of the deviation, is

$$S(N,k) = \sum_{j=M(N,k\text{-tuplets})}^{N} \binom{N}{j}(1/T^k)(1-1/T^k)^{N-j} \qquad (1)$$

This expression does not appear explicitly in Tart and Dronek (1982) because they illustrate their algorithm via an example, but the example is really illustrative enough to be sure that they use (1) or some equivalent, more economically computable expression.

The smaller is S(N,k) the more significant is the departure from serial independence in the set of those k-tuplets whose first k-1 members are equal to the actual k-1 target numbers preceding the one to be predicted, according to TD. Therefore it is logical to use for prediction the information provided by those k-tuplets for which the departure is largest i.e. $S(N,k)$ is smallest. Let this k is k_o, i.e.

$$S(N,k_o) = \min_{(k)} S(N,k)$$

then the prediction for the (N+1)th target is the X_o which occurred up to now most frequently after the last actual k_o-1 targets. This is the X_o, of course, for which

$$C(t_1,t_2,\ldots,t_{k_o-1},X_o) = \max_{(X)} C(t_1,t_2,\ldots,t_{k_o-1},X)$$

where $t_1,t_2,\ldots,t_{k_o-1}$ are the last k_o-1 targets, t_{k_o-1} being the Nth one. In the case of a tie the program makes a random decision between the equally best candidates produced by this procedure.

TD applied this algorithm to sequences of digits generated by an electronic RNG and to sequences of digits guessed by experimental subjects (Tart, 1976). The number of trials in these sequences was about 500 and the number of targets was 10, the digits 0,1,...,9. As it was expected the human-made sequences were highly predictable by the PPP. For a few RNG sequences the program also was able to produce significant ($p<.05$) hitting but its performance was clearly inferior to that of the best actual subjects of the experiment.

Here the natural question arises if there can be some more powerful inference strategies than used by the PPP. Tart and Dronek (1982) themselves mention this possibility in their paper. The strength of a stochastic predictor algorithm, of course, can not be defined for all possible data sequences or all random target generators. For example, if a given algorithm gives its prediction for the Nth target by computing a function $t_N = f(t_1,\ldots,t_{N-1})$ of previous targets $t_1,t_2,\ldots,t_{N-1}$ one can always construct a particular (biased) pseudo-random target generator so that its Nth target after $t_1,t_2,\ldots,t_{N-1}$ is any but not equal to $f(t_1,t_2,\ldots,t_{N-1})$. For this generator the algorithm at hand will never produce even a single hit. Evaluating a predictor algorithm or comparing some of them can only be done with respect to well-defined classes of generators i.e. to generators with well-defined types of bias.

The probabilistic predictor of TD aims at predicting targets coming from a generator with stationary, singlet-level or sequential biases. Stationary means that the conditional probabilities of targets following any given sequence are not depending on the absolute position in the whole sequence of targets ('time'):

$$p\{t_N | t_{N-s}, t_{N-s+1}, \ldots, t_{N-1}\} = p\{t_{N+M} | t_{N+M-s}, t_{N+M-s+1}, \ldots, t_{N+M-1}\}$$

for any N, M, and ,s where all t's can stand for any possible target.

Sequential bias means that these conditional probabilities (for some or all s) are not equal to 1/T where T is again the number of possible targets, as it were if all sub-sequences had equal chance to occur. Singlet-level bias means, of course, that the probability of occurrence of the single targets are not equally 1/T.

It is very likely that just these types of bias are most relevant in connection to parapsychological experiments. For me it is also very likely that the philosophy of the TD algorithm is basically correct. In fact I have already applied the PPP and it indeed pointed out some biases in a commercial RNG. One point in the TD algorithm is, however, arbitrary, and for a rather large class of biased generators leads to clearly suboptimal results. First I shall illustrate this by an example.

AN EXAMPLE OF NOT FULLY ADEQUATE OPERATION OF THE PPP

Let the targets be numbers 1, 2, and 3, and let the primary, stationary bias in the operation of the RNG be defined at the level of doublets, i.e. two-element sequences. Mathematically it means that the sequence of random numbers is Markovian and homogeneous. The conditional probabilities are:

$$p\{1|1\} = p, \quad p\{2|1\} = 1-p, \quad p\{3|1\} = p\{1|2\} = p\{2|2\} = p\{2|3\} = p\{3|3\} = 0, \quad p\{3|2\} = p\{1|3\} = 1$$

In the usual matrix form the matrix of transition probabilities is

$$A = \begin{bmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

This generator, of course, is very strongly biased and testing of it does not require sophisticated algorithms; but for illustration it is more important that it can be described analytically in a simple way and it will uncover just that weakness in PPP which can be manifested for more realistic generators too.

The bias on the doublet level induces bias on the singlet level as well. The singlet-level probabilities at the early stages depend on the actually first target but as the number N of trials becomes large they will approach their stationary values. These values can be determined from the condition

$$A^n \cdot A = A^n$$

that means roughly that the n-step transition probabilities for large N become independent of N; then the singlet-level probabilities are

$$p\{1\} = p^{(N)}\{1|1\} = p^{(N)}\{1|2\} = p^{(N)}\{1|3\}$$

$$p\{2\} = p^{(N)}\{2|1\} = p^{(N)}\{2|2\} = p^{(N)}\{2|3\}$$

$$p\{3\} = p^{(N)}\{3|1\} = p^{(N)}\{3|2\} = p^{(N)}\{3|3\}$$

In the example the stationary condition gives a system of equations and the solutions are

$$p\{1\} = 1/(3-2p), \quad p\{2\} = p\{3\} = (1-p)/(3-2p) \qquad (2)$$

What will the PPP do for the various targets of this generator if it must choose between decisions based on singlet- and doublet-level considerations? Assume that the Nth target was 2. Then from the matrix A the next, (N+1)th one must be 3 as $p\{3|2\}=1$. So the counters C(2,1) and C(2,2) will show always zero and the doublet-level prediction will always be 3. The singlet-level prediction depends on the actual random transitions following target 1 but if $p>0$ the prediction on the average will be 1 because then $p\{1\}>p\{2\}=p\{3\}$. If the actual content of counter C(1) deviates more significantly from its unbiased value N/3 than counter C(2,3) does so from its unbiased value N/9 (in fact (N-1)/9 but the difference for large N is negligible) then the ultimate prediction of the PPP will be 1, a miss. The statistical significance for large N can be determined by approximating the binomial distribution with a normal (Gaussian) one on both the singlet and the doublet levels. The p-value of significance for singlets is smaller than that for doublets if and only if the corresponding Z-score for singlets is larger than that for doublets because for normal distribution there is a one-to-one relation between Z-score and significance level. Now for singlets in our case

$$Z(1) = \frac{C(1) - N/3}{\sqrt{[N.(1/3).(2/3)]}} \qquad (3)$$

and for doublets

$$Z(2) = \frac{C(2,3) - N/9}{\sqrt{[N.(1/9).(8/9)]}} \qquad (4)$$

Now let us assume that C(1) and C(2,3) are just equal to their respective expected values given the transition matrix A and the probabilities derived from it. This assumption will not be valid in all cases, of course, but it will show the relevant relation between the Z-scores on the average. So

$$C(1) = N \cdot \frac{1}{3-2p}$$

$$C(2,3) = N \cdot p\{2,3\} = N \cdot p\{2\} \cdot p\{3|2\} = N \cdot \frac{1-p}{3-2p}$$

Then the ratio R_{12} between expected Z(1) and Z(2) will be

$$R_{12} = \frac{E[Z(1)]}{E[Z(2)]} = \frac{[N.\ \ 1/(3-2p)-N.(1/3)] \cdot \sqrt{[N.(1/9).(8/9)]}}{[N.(1-p)/(3-2p)-N.(1/9)] \cdot \sqrt{[N.(1/3).(2/3)]}}$$

$$= \frac{4p}{18-21p}$$

Missing occurs regularly if $R_{12}>1$ from which $p>18/25=.72$. Note that R_{12} does not depend on N so the PPP can not 'learn' the sequence 2,3 with increasing number of trials if $p>18/25$. This happens in a situation where target 2 is always followed by target 3, a most marked kind of bias that can reasonably expect to be recognized by a predictor algorithm.

IMPROVEMENT OF THE PPP

What is the origin of the failure of the PPP in recognising the deterministic sequence 2,3 in the example? By doublet-level decision it would recognize it. The problem is that if p and consequently the singlet-level frequency of target 1 is large then the statistical significance of 1 outweighs that of (2,3) and the program logically chooses to decide according to the singlet level. If p, the conditional probability $p\{1,1\}$, is larger than a threshold value, on the doublet level there are too many (1,1) doublets and too few (2,3) ones in spite of the fact that among (2,X) doublets (2,3) is relatively the far most frequent. But in PPP the decision between levels is based on the absolute statistical significance of doublets, triplets, etc., not realizing that the prediction after a particular $t_1,t_2,\ldots,t_{k-1}$ sequence (on the level of general k-tuplets) has no use of the information coming from k-tuplets with first k-1 elements other than $t_1,t_2,\ldots,t_{k-1}$. Really the relevant deviation here is not that of the largest $C(t_1,t_2,\ldots,t_{k-1},X)$ from its unbiased, overall expected value $1/T^k$ but its unbiased, conditional expected value given the first k-1 known elements $t_1,t_2,\ldots,t_{k-1}$.

The number of all previous k-tuplets with first k-1 members $t_1,t_2,\ldots t_{k-1}$ is

$$S(t_1,t_2,\ldots t_{k-1}) = \sum_{X=1}^{T} C(t_1,t_2,\ldots t_{k-1},X).$$

If all targets X were equally probable to follow this first k-1 targets then the expected values of the ratios

$$C(t_1,t_2,\ldots t_{k-1},X)/S(t_1,t_2,\ldots t_{k-1})$$

would be all equal to 1/T. This ratio does not depend on k, therefore the deviations of the numbers of most frequent k-tuplets from their expected conditional numbers can be compared directly without any significance test. This yields a computationally simpler algorithm than the PPP. All we have to do is to find the largest $C(t_1,t_2,...t_{k-1},X)$ for every k (just as in PPP), compute the ratio of its content to the corresponding sum of all $C(t_1,t_2,...t_{k-1},X)$ and choose the one for which this ratio is the largest.

An example: Let the number T of possible targets be 5; the k-tuplets considered ('maximal memory length' in the notation of TD) are up to triplets. Let the last input target be 4, and the one before it 1, and let the contents of the relevant counters be:

C(1) = 1004	C(2) = 983	C(3) = 1011	C(4) = 996
C(5) = 1006	C(4,1) = 102	C(4,2) = 98	
C(4,3) = 96	C(4,4) = 98	C(4,5) = 100	C(1,4,1) = 21
C(1,4,2) = 16	C(1,4,3) = 21	C(1,4,4) = 20	C(1,4,5) = 22

The sum on singlet level is 5000, on doublet level 494, and on triplet level 100. The largest ratio on singlet level is for target 3, its value is 1011/5000=.2022; on the doublet level it is for target 1, its value is .2065; on the triplet level it is for target 5, its value is .22. The program finds the largest ratio on triplet level and chooses accordingly target 5 as its prediction.

Note that in this way the decision does not depend on the absolute frequency of triplets (1,4,5), but only its relative frequency among triplets (1,4,X) so the problem which PPP encountered with in the example of the previous chapter is avoided. For a large number of trials these ratios approximate well just the conditional probabilities of the targets that are relevant in characterising the biases, therefore with increasing N this method aproaches an optimum.

EXPERIMENTAL COMPARISONS BETWEEN PPP AND PPP-B

Simulation of biased random number generators.

All computers can generate pseudo-random fractional numbers between 0 and 1, all with equal probabilities. If one wants a particular target to occur with probability p the simplest method is to ask the value of the next pseudo-random number R and decide that the next target will be t if and only if R is less than or equal to p. In BASIC code:

```
R=RND: IF R<=P THEN TARGET=T
```

If there are more targets the interval [0,1) must be divided in proportions corresponding to the required probabilities of the targets and then the decision strategy is similar. For example, if the targets are the numbers 1, 2, and 3 and the required probabilities are $p\{1\}=.2$, $p\{2\}=.35$, $p\{3\}=.45$ then the corresponding BASIC code is

```
R=RND: IF R<=.2 THEN TARGET=1 ELSE IF R<=.55
THEN TARGET=2 ELSE TARGET=3
```

If the probabilities depend on some previous targets, as in cases of sequential biases, the same technique applies for each previous target or target combination, with writing one conditional statement for each. For example if the targets are again 1, 2, 3 and the transition matrix is that of the chapter 'the example of not fully adequate operation of the PPP' then the code is

```
R=RND
IF TARGET(N-1)=1 THEN IF R<=P THEN TARGET(N)=1 ELSE
TARGET(N)=2
IF TARGET(N-1)=2 THEN TARGET(N)=3
IF TARGET(N-1)=3 THEN TARGET(N)=1
```

In this way sequential bias can be simulated.

The effect of maximal memory length

In their experimental investigations with PPP Tart and Dronek tested sequential biases up to the sextuplet level. They claimed that the power of the predictor program increases in principle with increasing number of levels it uses. For infinitely long sequences it is surely true but for finite ones the usefulness of a given level depends heavily on the number of targets and the series length. Let us see again a simple example.

Let the number of possible targets be 10 as it was in the TD study and let the sequence length be 500. Consider the level of sextuplets. From 10 targets 10^6 sextuplets can be constructed because any target on the first place can be followed by any one on the second place (10x10 possibilities); any of these doublets can be followed again by any target on the third place (10x10x10) and so on up to 10^6 . If the probabilities of all targets are approximately the same, all individual sextuplets have the probability of occurrence of about 10^{-6}. Hence the probability that a particular sextuplet occurs twice or more in 500 trials is (by the binomial distribution) roughly 1.3×10^{-7} that is practically zero. If there is some bias on the sextuplet level, e.g. one which doubles the probabilities of some sextuplets and halves those of others it has no chance of manifestation in 500 trials. Even totally deterministic biases will remain hidden because all this is valid for quintuplets as well and there will be no quintuplets for the determined sixth target to follow them. Even biased triplets have little chance of being uncovered; the a priori probability of two or more occurrences of any particular triplet in 500 trials is still about .09.

For this reason in the following simulation I shall use singlets, doublets and triplets only, with targets 1, 2, and 3. As it can be seen readily, the smaller the number T of targets, the more levels can be used realistically with a given number of trials which in the following will always be 500.

Symmetrical bias on doublet level

Symmetrical bias means that the increased conditional probabilities of a given target after some subsequences are compensated by decreased ones after some other subsequences so that on lower levels there will be no bias at all. Let the transition matrix of the biased generator be:

$$A_1(B) = \begin{bmatrix} 1/3-B & 1/3+2B & 1/3-B \\ 1/3 & 1/3 & 1/3 \\ 1/3+B & 1/3-2B & 1/3+B \end{bmatrix}$$

where $0 \le B \le 1/6$. B is the amount of bias that can be varied. In the simulation experiments I have used B=.05. As the singlet-level probabilities now are all equal to 1/3 it is probable in advance that the failures predicted for PPP will not occur in this case. Table 1 shows the scores produced by PPP and PPP-B for maximal memory lengths 2 and 3. Using only singlet level is not adequate for comparison because on singlet level both algorithms predict always the same target.

TABLE 1
PPP and PPP-B number of hits in 500 trials with three targets in the case of symmetric bias on the doublet level for different length of memory

Max. memory length	2				3		
Algorithm	PPP	PPP-B	D		PPP	PPP-B	D
Exp. 1	193	197	4	Exp. 6	179	171	-8
2	197	195	-2	7	182	182	0
3	172	179	7	8	185	179	-6
4	190	191	1	9	185	194	9
5	182	192	10	10	188	190	2
Average	186.8	191.8	4.0		183.8	183.2	.6
Sq. root of variance for D=			4.2				6.8

NOTE: In each experiment the input sequence was the same for PPP and PPP-B pairwise, therefore for comparison a one-sample t-test is adequate. The results of this comparison are not significant. The expected value of average hits assuming only chance effect (null hypothesis) =166.7; scatter of this average by null hypothesis is 4.7. All averages are separately significant at least at the .003 level. The difference D=hits(PPP-B)-hits(PPP).

The simulation results can be summarised for symmetric bias of the tested type as follows:

1 - Both algorithms produced significantly more hits than chance expectation.
2 - There were no significant differences between the scores of the two algorithms.
3 - Inclusion of triplet level did not improve the performance for

either of the two algorithms.

So we can say that for symmetrical bias of this kind (and very probable for other kinds as well) PPP-B is not weaker than PPP and both programs perform as expected. The advantage of PPP-B for such biases is only its relative simplicity and greater speed.

Asymmetrical bias on the doublet level

Let the transition matrix be:

$$A_2(B) = \begin{bmatrix} .5 & .3 & .2 \\ .2 & .3 & .5 \\ .5 & .3 & .2 \end{bmatrix}$$

From the stationary condition the singlet-level probabilities are:
p{1}=.41 p{2}=.3 p{3}=.29
So we can expect some misses on the part of the PPP in sequences 2->3. The results of simulation are shown in Table 2, again for maximal memeory length 2 and 3.
The main points are:

1. Both algorithms produced significantly more hits than chance expectation.
2. PPP-B produced significantly more hits than PPP.
3. Inclusion of triplet level did not improve significantly the performance for either algorithms (t-test with 4 degrees of freedom) but for both ones it tended to do so slightly. The reason is that the bias on doublet level naturally appears also on triplet level and helps to overweigh initial transients in the singlet frequencies.

Asymmetrical bias on the triplet level.

The transition matrix is the same as the preceeding one but the probabilities of targets do not depend on the previous target but on the one before it i.e. two targets back. In this way there is no bias on the doublet level other than that which follows from the induced singlet-level one. The result is that with maximal memory length 2 both programs predict always target 1, after, of course, some initial

TABLE 2
PPP and PPP-B number of hits in 500 trials with three targets in the case of asymmetrical bias on the doublet level for different length of memory

Max. memory length	2				3		
Algorithm	PPP	PPP-B	D		PPP	PPP-B	D
Exp. 11	200	219	19	Exp. 16	235	252	17
12	227	245	18	17	232	248	16
13	216	248	32	18	228	260	32
14	219	240	21	19	230	252	22
15	202	224	22	20	212	238	26
Average	212.8	235.2	22.4		227.4	250.0	22.6
Sq. root of variance for D=5.6							6.6

Notations as in Table 1. All averages are again significant as above, and here also the differences are significant at .05 level.

fluctuations. So for comparison it is no use to consider maximal memory length 2. With length 3 the results are shown in Table 3. They indicate again what we have seen in the case of asymmetric bias on doublet level: PPP-B was again stronger as it was expected.

CONCLUSIONS

In their paper Tart and Dronek (1982) propose that "The PPP (or some superior version of it) be adopted as the standard measure of predictability, and used in all psi experiments in which percipients receive feedback on target identity before the end of the experiment". In my opinion this proposal is fairly well founded and as the new

TABLE 3
PPP and PPP-B number of hits in 500 trials with three targets in the case of asymmetrical bias on the triplet level

Algorithm		PPP	PPP-B	D
Experiment	21	206	215	9
	22	206	226	20
	23	198	220	22
	24	187	204	17
	25	221	245	24
Average		203.6	222.0	18.4
Sq. root of variance for D = 5.9				

Notations as in Table 1. Maximal memory length is 3. All averages are again significant as above, and the difference is so at the .05 level.

version PPP-B is indeed superior slightly in some situations and not inferior in others, and as it is simpler to implement on microcomputers, I recommend this PPP-B for the purpose outlined above.

As a last remark, regarding to the findings of TD on experimental sequences in C.T.Tart's training study, it is quite obvious that the very large discrepancy between the scores of the PPP and those of the best percipients could not be decreased dramatically by using the PPP-B. In sequences of as few as 500 trials simply there is no room enough to accumulate differential scores for this purpose. Therefore their assertion that in the investigated training study "ESP is the best way to account for the hitting" is not weakened essentially by the existence of this better version of the PPP. This point should be kept in mind in particular regarding to Tart's later analysis on temporal displacement effects (Tart, 1979) which may be very important in seeking consistent patterns in psi test results.

ABSTRACT

The Probabilistic Predictor Program (PPP) is a useful tool for uncovering regularities in apparently random target sequences but for some types of bias in sequential probabilities its performance is clearly suboptimal. An example is given to illustrate a typical situation with excess misses on the part of the program. The origin of excess misses is the use of unconditional frequencies of subsequences instead of more natural conditional ones in the choice process between decision levels. A new version PPP-B of this algorithm is proposed which corrects the above inadequacy. Simulation tests show that PPP-B produces significantly more hits than PPP for sequences coming from asymmetrically biased generators and that PPP-B is equivalent to PPP for symmetrically biased ones. The new version is simpler to implement especially on microcomputers; a complete BASIC code list is given.

REFERENCES

Tart, C.T. 'Learning to use extrasensory perception', Chicago, University of Chicago Press, 1976.

Tart, C.T. 'Improving real-time ESP by suppressing the future: trans-temporal inhibition', in C.T. Tart, H.E. Puthoff, R. Targ, Preager (eds.) 'Mind at large', New York, 1979.

Tart, C.T., Dronek, E. 'Mathematical inference strategies versus psi: initial explorations with the probabilistic predictor program', E.J.P., 4, 1982, 325-355.

APPENDIX

BASIC code for PPP-B. This is a rather common BASIC dialect although some modifications may be needed in implementation. The names of variables are either self-explaining or mentioned in the text.

```
10  DEFINT C,I,J,K,N,T,S
20  INPUT NMAX,T,KMAX
30  DIM TARGET(NMAX): DIM PREDICTION(NMAX): DIM COUNTER(KMAX,
    T*KMAX)
40  HITS=0: GOSUB 300: PREDICTION(0)=INT(RND*T)+1
50  FOR N=1 TO NMAX: GOSUB 340
60  IF TARGET(N)=PREDICTION(N-1) THEN HITS=HITS+1
70  FOR K=1 TO KMAX
80  IF N<K THEN GOTO 100
90  I=0: FOR I1=0 TO K-1: I=I+TARGET(N-I1)-1)*T**I1: NEXT I1:
    COUNTER(K,I)=COUNTER(K,I)+1
100 NEXT K: MAXRAT=0
110 IF N<KMAX THEN PREDICTION(N)=INT(RND*T)+1: GOTO 220
120 FOR K=1 TO KMAX
130 I=0: FOR I1=0 TO K-2: I=I+(TARGET(N-I1)-1)*T**(I1+1):
    NEXT I1
140 SUM=0
150 FOR I1=0 TO T-1
160 SUM=SUM+COUNTER(K,I+I1): NEXT I1
170 IF SUM=0 THEN GOTO 210
180 FOR I1=1 TO T: RATIO=COUNTER(K,I+I1-1)/SUM
190 IF RATIO>MAXRAT THEN MAXRAT=RATIO: PREDICTION(N)=I1
200 NEXT I1
210 NEXT K
220 NEXT N: PRINT "HITS=";HITS: END
300 REM RANDOMIZATION
    This subroutine depends on the actual number generating
    procedure of the computer used and accordingly the
    BASIC statements for this purpose.
 .. RETURN
340 REM INPUT TARGETS
    This subroutine depends on the random number generator
    tested and on the interface with it.
 .. TARGET(N)=...
 .. RETURN
```

Zoltan Vassy,
1085 Budapest,
Jozsef krt. 49. I. 6/B,
Hungary.

NOTE, IN ADDITION TO THE ARTICLE BY Z. VASSY

An implicit assumption in Mr. Vassy's expose is that the subject is naive about the random number generator before the first trial. However, if one assumes that the subject did T training trials before his N experimental trials, one should reckon with the performance of the predictor program on the last N out of a total of T+N trials. Mr. Vassy's program can easily be extended in this way.

The consequence for the evaluation of C.T. Tart's experiments is that the number of training trials his subjects performed before the experiment proper is not irrelevant.

Joop M. Houtkooper,
Jan Swammerdam Institute,
University of Amsterdam,
P.O. Box 60.000,
1005 GA Amsterdam,
The Netherlands.

EUR.J.PARAPSYCHOL., V.5, 1984

APPLYING THE 'WORRY' HYPOTHESIS TO SPONTANEOUS PARANORMAL EXPERIENCES

Sybo A. Schouten
University of Utrecht

In any field of science experimental research is to a certain extent based on some unproven assumptions and paradigms. From the parapsychological literature it can be inferred that most experimental research in this field is based on the assumption that psi exists and is to be considered as a kind of human faculty. Thus subjects are employed in experiments as if they are able to apply psi 'abilities' to receive telepathic impressions or to exert psychokinetic influence on a random number generator.

Although experimental evidence is available to support the assumption of the existence of psi, hardly anything is known about its nature. Partly because of this lack of understanding the concept of psi can only be defined in a negative way. As a consequence psi is considered to explain experimental results only when all other possible explanations have to be rejected.

Because of the negative nature of the definition of psi the concept of psi can hardly be put to an efficient test of refutation. To falsify an hypothesis seems much more difficult than to falsify the results of an experimental procedure. If it is predicted that in a PK-experiment a significant deviation from chance expectation will be observed, one test suffices to decide whether or not such significant deviation is observed. However, to falsify the concept of psi as an explanatory agent is more complicated. If in a properly designed and executed psi experiment significant results are obtained the operating principle is assumed to be paranormal. However, if the results are in accordance with the expected probability distributions, then it is concluded that under the experimental conditions the subjects were not able to use their psi abilities. Thus, unless it can be proven that an alternative explanation is applicable to the results of that particular experiment, the concept of psi as the operating principle cannot be falsified by the experimental results. This "low degree of falsification" and the associated lack of repeatable experiments, which would allow us to demonstrate at least the effects of the assumed process at will, can be considered the main reasons for the fact that the scientific merits of this type of research are still in debate.

A different approach to research in this field focuses on spontaneous paranormal experiences (SPES). SPES are occurrences in the daily life of people which usually involve a correspondence between the content of a subjective experience (for instance a dream) and an event in which the correspondence impresses people as meaningful, not coincidental, and not explicable by ordinary laws of nature and therefore as suggestive of psi. Thus SPES are a class of subjective experiences to which a psi character is attributed. Such experiences are still frequently reported and their occurrence is probably an important factor in the widespread belief in psi. In fact, it could be argued that it is questionable whether experimental research in this field would have continued to the present if people had not reported SPES on the scale they have.

An important advantage of the study of SPES is that it focuses on the experience itself as the primary research object. The study of human experiences is quite common in psychology and generally does not require any specific presuppositions. Hence this approach in parapsychological research does not require the a priori assumption of the existence of psi in order for the research to become meaningful. The characteristics of the experiences can be studied and various hypotheses (parapsychological, psychological, physiological, etc.) can be applied to the material to study to what extent they are able to explain the characteristics observed.

Thus although SPES are suggestive of psi the study of these experiences allows a more falsificationistic approach. It is possible to test the applicability of various hypotheses of a non-psi nature and the outcomes of such studies could lead to the conclusion that the concept of psi as an explanation for this class of experiences has to be rejected. On the other hand, with this approach the concept of psi could gain a much stronger position if it stands up to such attempts at falsification and turns out to survive the competition with other hypotheses.

Based on this approach a series of studies on SPES have been carried out. The aim of the first study (Schouten, 1979) was to investigate whether these experiences show common characteristics and if meaningful relationships between variables associated with these experiences could be found. It was hypothesized that if SPES are not based on psi but merely constitute a collection of random occurrences to which a paranormal character is attributed, then few or no common characteristics or meaningful relationships could be expected.

From this study which involved quantitative analyses of a collection of SPES it was concluded that these experiences can be considered as a distinct class of experiences and not as a collection of random events. Also in this study a number of non-psi hypotheses, so-called reporting effects, were tested to investigate whether they could explain some of the observed characteristics. This turned out not to be the case.

The data of two subsequent studies (Schouten, 1981, 1982) based on different collections of SPES confirmed to a large extent the findings of the first study. A further objective of these confirmatory studies was to test the applicability of another non-psi hypothesis: the 'cultural' hypothesis. If a non-psi hypothesis is applied to these experiences then the hypothesis should explain why a large number of people single out one or a few events in their lives which they label 'psi' in such a way that together these events show the characteristics as observed in these studies. The 'cultural' hypothesis assumes that in a given culture people have a specific image of what constitutes a paranormal experience and that only those random events in their lives which fit that image are labelled 'paranormal'. This could explain how a collection of only random events still could result in a collection of ostensible paranormal experiences with common characteristics.

But if this hypothesis is valid differences in characteristics of these experiences are to be expected when collections are compared from cultures with different 'images' of a paranormal experience.

Hence the confirmatory studies mentioned above were based on collections of SPES from different countries and epochs.

The results of the two confirmatory studies yielded little evidence for the validity of the 'cultural' hypothesis. In most instances the characteristics of the experiences in the different collections and the relationships between variables appeared to be the same. The only variable of importance which seems to be influenced by cultural differences is what L.E. Rhine called the 'form of experience', the way the subjective experience becomes conscious.

The finding that the results of the three studies confirm each other so well is all the more remarkable because there is reason to assume that the different collections of SPES were acquired from different samples of the population. However, if the characteristics of these experiences are largely independent of differences in sampling, culture, and epoch, then it follows that it is justifiable to use data from present-day studies for comparison with the data obtained in the studies discussed above.

Comparing data on certain characteristics of present-day samples with the known characteristics of SPES might be useful in gaining a preliminary indication of the merit of a potential explanatory hypothesis. Too much difference between the rather stable characteristics of paranormal experiences and the data associated with a particular hypothesis is an indication that more extensive research into that direction is of little promise. This strategy enables us to screen a number of potential hypotheses with relatively little research and to concentrate the more elaborate research of the manipulative kind on those hypotheses which have passed this screening.

Based on this strategy two studies were carried out to test another hypothesis of a non-psi character: the 'worry' hypothesis. This hypothesis relates SPES with the tendency of people to worry about possible future negative events. Most people realize that in many instances the probability is rather low that the events they fear will materialize. Hence it could be that a paranormal character is attributed to such feelings of worry in those cases in which, despite the low probability, the events they worried about do happen. For instance, it might be that parents who worry about a possible accident each time their child is out on the streets attribute a paranormal character to their feelings when the child unexpectedly does become the victim of an accident.

If SPES are a consequence of the human tendency to worry then some predictions can be made with regard to relationships between characteristics of the tendency to worry and of SPES. Considering the tendency to worry one can distinguish three main aspects: (1) The possible events one worries about. (2) The persons one worries about (including onself). (3) The extent to which one worries about persons and about possible events. Each of these aspects result in one or more predictions.

The possible events people worry about can be expected to be approximately distributed as the events involved in SPES. The same applies to the persons one worries about. Because females compared to males more often report SPES, according to the hypothesis one would expect to find that females also tend to worry more frequently than males. In addition, since in most SPES a different person is involved one should also expect to find that people tend to worry more often about other persons than about themselves.

A literature survey did not yield useful data on the everyday worries of common people. Hence two studies were carried out to acquire data on the various aspects of the tendency to worry and to compare these data with the known characteristics of SPES.

These studies are presented here. The two studies involved different samples of subjects. In both studies the variables 'event' (what does one worry about) and 'person' (which person does one worry about) were measured but with different methods. Because in the two studies the data were acquired with different methods and from different samples an estimation of the variation in the distributions of these variables is obtained. This provides an indication whether the associated distributions observed for SPES are more or less within that range or clearly outside that range. One of these studies, the second one described below, was also used to develop scales for measuring the relative strength of the extent each respondent worries about him- or herself and about other persons.

FIRST STUDY

METHOD

Because it was assumed that respondents might consider their worries too intimate to discuss their exact nature, they were requested only to classify some aspects of the worries they had at the time of participating in this study. These aspects involved the variables 'event' and 'person'. For each of these variables categories were devised in such a way that the obtained data would be comparable to the data obtained in the studies on SPES.

As regards the variable 'event' subjects were requested to classify the nature of their worry or problem using the response categories (1) death, (2) serious illness or accident, (3) light illness or accident, (4) serious financial or material matters, (5) slight financial or material matters, (6) worries related to work, school, or daily matters, (7) psychological problems, (8) worries related to the general (political) situation, and (9) worries of another nature. As regards the variable 'person' the applied categories were (1) yourself, (2) partner, (3) member of the family, (4) some other relative, (5) friend, (6) acquaintance, (7) more than one person. In addition data were obtained concerning sex and age of the respondent. Age was indicated in categories of 10 years (thus 10-20, etc.).

Subjects were requested to fill in one questionnaire for each worry or problem they had at the time of participating in this study starting with the problem which occupied them most at the moment. The sample size was specified prior to collecting the data to be approximately 100.

RESULTS

Questionnaire data were obtained from 101 subjects of which 51 were males and 49 were females (one respondent failed to answer this question). Of the subjects 51 were below the age of 30 and 50 above that age. The average age of the subjects was approximately 37 years.

The 101 subjects filled in questionnaires for 208 worries or problems. In only 3 instances did respondents use the category 'worries of another nature' to indicate the nature of their worry. Hence it can be concluded that the categories used covered the nature of the worries of the respondents quite well.

Each subject filled in at least one response sheet. Because subjects were requested to start with the worry which occupied them most at the time, it is possible to distinguish between 'most dominant worries' (101) and 'less dominant worries' (107).

Table 1 presents the distributions obtained for the different types of possible events the subjects worried about. In this table the categories 'death' and 'serious illness or accident' are combined into one category. The same applies to the categories 'serious financial or material matters' and 'slight financial or material matters'.

TABLE 1
Distribution of possible events respondents worry about for most dominant worries and less dominant worries

	death serious accidents	light illness accidents	material matters	daily matters	psychol. problems	general situation
most dominant	15	4	8	39	20	14
less dominant	19	10	19	18	25	14
total	34	14	27	57	45	28
in %	16%	7%	13%	28%	22%	14%

The distribution of events for most dominant worries and less dominant worries apear to differ to a significant degree (Chi-square=15.7, df=5, $p<.01$). In the case of most dominant worries problems concerning daily matters, work, or school rank highest, followed by psychological problems and worries related to serious threats to health. In the case of less dominant worries problems of a psychological nature rank highest, followed by material worries and problems concerning serious threats to the health. In both distributions problems concerning slight illnesses or accidents rank lowest.

The distributions for most dominant worries and less dominant worries as regards the person about which the respondents worry are presented in table 2. In this table the categories 'partner' and 'member of the family' are combined. The same holds for the categories 'friends' and 'acquaintances'.

TABLE 2
Distribution of persons respondents worry about for most dominant and less dominant worries

	self	family	other relatives	friends acquaintances	more than one person
most dominant	44	27	6	6	18
less dominant	37	29	9	9	23
total	81	56	15	15	41
in %	39%	27%	7%	7%	20%

The two distributions appear not to differ (Chi-square=2.31, df=4). Combining the data from the tables 1 and 2 it appears that the prominence of a worry is in general more dependent on the nature of the problem than on the person involved. From the data of table 2 it also follows that most worries concern either the respondent him- or herself or the family in which he or she lives. As regards the category 'more than one person' it should be noted that the majority of the worries classified in this category concern general problems, because worries about the general situation are always scored as being related to more than one person.

The frequency matrix of the variables 'event' and 'person' combined contains 30 cells, which for a total of 205 problems result in too many cells with an expected frequency below 5 to make a chi-square analysis meaningful. However, a global inspection of the data does not indicate strong discrepancies between observed and expected values. The most striking deviations suggest that worries concerning daily matters, work, and school, and worries concerning material matters are more frequently related to the respondent him- or herself whereas worries about serious threats to the health and worries of a psychological nature mostly concern relatives not belonging to the own family.

Male and female respondents appear to be evenly distributed over the various age categories. No differences are observed between the two age groups under and above 30 years of age as regards the number of worries for which response sheets were filled in as well as regards the nature of the problems. However, a slight difference appears when the two age groups are compared as regards the distribution of persons involved in their worries. Younger people tend to worry more often about themselves or about friends whereas older people worry relatively more frequently about their family and relatives (Chi-square=9.70, df=4, $p<.05$).

Males and females report on the average an equal number of worries. No relationship is observed between sex of respondent and the variable 'event'. A slight difference between males and females is observed in the distributions of the persons they worry about. Males tend to worry more often about themselves whereas females worry relatively more often about other persons (Chi-square=11.4, df=4, $p<.05$).

SECOND STUDY

This study was carried out with two aims. One aim was to obtain data on the same variables as studied in Study I but with a different method and with different samples. In this respect the study can be considered as a confirmatory study. Another aim was to develop scales to measure the extent to which each person worried about him- or herself and about others.

The data collected in the first study were based on all problems for which a response sheet was filled in. Therefore all analyses discussed so far are exclusively based on number of worries. Thus out of necessity all worries have been treated as if they were of equal importance to the respondents. The only indication of differences in importance is given by the rankorder applied by the respondents who have filled in more than one response sheet. However, conclusions from these data have to be treated with caution because it can be assumed that extensive differences as regards seriousness of problems exist between subjects. What we may conclude from the first study, though, is that the number of worries mentioned related to other persons is higher than the number of worries mentioned related to the respondent him- or herself (61% versus 39%).

The scale construction attempted in the second study was aimed at obtaining scales which would enable one to compare the respondents on the same scales as regards the extent to which they worry about themselves and others. Of course such scales can only be rankorder scales. Nevertheless, if such scales could be obtained it would provide an indication about the relative strength people tend to worry about themselves and about others and about the relationship between these two variables. In other words, do people worry more strongly about others than about themselves and do the data justify one to distinguish between 'worriers' and carefree people?

METHOD

Subjects were presented with a questionnaire which consisted of 39 statements concerning the tendency to worry about themselves and others. These statements were followed with two 'open' questions. In the first one the respondent was asked to describe the nature of the worries he or she had at the time. The second question concerned the persons the respondent worried most about. Subjects were requested to indicate the relationship with these persons, for instance, spouse, child, friend. In addition data were collected about age and gender of the respondents.

RESULTS

Exactly 100 respondents, 48 males and 52 females, returned usable questionnaires. Of the respondents 64 were under 30 years of age and 36 above 30 years of age. The average age of the respondents was approximately 37 years. Males and females appear to be equally distributed over the various age groups and males and females appear not to differ in the average number of worries they describe.

Event and person of worry

In order to obtain data comparable to those acquired from the first study the analyses involved only descriptions of worries and problems in which both the nature of the worry and the person it concerned were given. Because of the way the descriptions were acquired in this study no distinction could be made between most dominant and less dominant worries.

The 100 respondents gave in total 170 complete descriptions. Based on these descriptions the experimenter rated each worry in the categories as applied in the first study. In table 3 the distribution of the nature of the worries is presented. Table 4 presents the distribution of the persons involved in the descriptions. Contrary to what was observed in the first study the data of table 4 reveal that in the present study the majority of the worries concern the respondent him- or herself. However, the rankorder over the various categories of persons appears not to differ from the rankorder observed in table 2.

An inspection of the frequency matrix of the variables 'event' and 'person' combined revealed that again the most striking (positive) deviation appears to be in the cell which combines the variable 'event'-category: 'daily matters, school, and work' with the variable 'person'-category: 'respondent him- or herself'.

In this study no relationship was observed between gender and nature of the problems described. However, in agreement with the findings of Study I a trend is noted that males tend to worry relatively more often about themselves whereas females tend to worry relatively more often about others. However, in this study the difference appears not

TABLE 3
Distribution of possible events respondents worry about

death serious accidents	light illness accidents	material matters	daily matters	psychol. problems	general situation
8	21	15	62	41	23
5%	12%	9%	36%	24%	14%

TABLE 4
Distribution of persons respondents worry about

self	family	other relatives	friends acquaintances	more than one person
96	27	4	15	25
57%	16%	2%	9%	15%

to reach a significant level.

Age appears to be significantly related to both the variables 'person' and 'event'. Like in the first study it was found that younger people (under 30 years of age) worry relatively more often about themselves and about friends whereas older people (over 30 years of age) worry more often about their family and relatives (Chi-square=24.2, df=4, $p<.01$). As regards the nature of the worries the difference appears to be that younger people worry more frequently about daily matters, work, and school, whereas older people mention relatively more often worries about serious health problems and about the general situation (Chi-square=32.7, df=5, $p<.01$).

The age-effects appear to be more pronounced in the data of this study than in the data of Study I. This might be due to the fact that compared to the previous study in the present study a greater number of younger respondents participated. Still the average age of the subjects does not differ for the two studies. This implies that the variance in the age distribution is higher in the present study and the higher variance might have contributed to the effects of age. The higher proportion of young respondents in this study will also have contributed to the finding of this study of a higher proportion of problems related to the respondent him- or herself (57% versus 39% in the first study) since from the data of both studies it follows that younger people tend to worry more often about their own affairs.

The scale construction

The questionnaire presented to the subjects included 39 statements related to various aspects of the tendency to worry about others and about oneself. These statements could be responded to by indicating 'yes' or 'no'. The aim of including these statements was to construct comparable scales to measure the extent to which respondents worry about themselves and about others. In order to obtain scales with comparable scale scores 15 related pairs of statements were constructed. The content of each statement of a pair was similar as regards the meaning of the statement. However, one statement of a pair referred to worries about the respondent him- or herself, whereas the other statement referred to similar worries about other persons. Although the meaning of each statement of a pair had to be the same the phrasing was made as different as possible to minimize the possibility that the response on one statement of a pair would influence the response on the other statement of the pair. For instance, two statements of a pair were: 'Do you occasionally worry about your future?' and 'Now and then the future of the people I care about worries me '.

For the same reason mentioned above, statements were presented in such an order that statements of each pair were separated by a number of other statements. Statements related to worries about the respondent him- or herself and statements related to worries about other persons were mixed in random order.

The 15 pairs of statements provided 30 statements. The 9 other statements included in the questionnaire were of a more general nature and supposed to reveal something of the respondent's inclination to worry. Examples of such statements are: 'Are you often feeling depressed?' and 'Did you have a lot to worry about lately?'.

A scale-construction technique developed by Mokken (1971; Kingma, 1984) was applied to the obtained data. This technique involves a multidimensional analysis which yield Guttman-type scales. A Guttman scale is based on the principle of consistency. For instance, if a pupil has to solve a number of problems of increasing difficulty which are all supposed to measure the same ability, then it should be expected that if the pupil solves a certain number of problems these problems are the ones which are supposed to be the least difficult. However, if pupils turn out to be able to solve the problems which are supposed to be the most difficult and not the problems supposed to be of average difficulty then the series of problems - the scale - is considered inconsistent. In that case something is wrong with the assumptions on which the degree of difficulty of the problems were based.

In an analogous way attitude statements can be tested for consistency. If a subject agrees with a statement which expresses a strong tendency to worry then it would be inconsistent if he or she also agrees with a statement which is supposed to indicate a relatively carefree attitude. Based on the responses of the subjects the applied scale analysis technique constructs scales made up of statements which are to a significant degree consistent with each other. It is a multi-dimensional technique because the analysis can result in more than one scale. In that case if the different scales include different statements the scales can be considered to reflect different attitudes. Although such scales are independent it is possible that a correlation is found between the scores of subjects on the two scales like, for instance, different sports may require entirely different skills but achievements may be still correlated.

The scale analysis applied to the responses of the subjects to the statements yielded two independent scales. Scale I includes 13 statements and Scale II includes 11 statements. The two scales have no statements in common. Scale I appears to consist of 7 statements related to worries about oneself, of 2 statements related to worries about others and of 4 statements of a general nature. Scale II includes 9 statements related to worries about others, 1 statement concerning problems related to the respondent and 1 statement of a general nature. The two scales were reduced to 10-item scales each by excluding the few statements which do not fit the general pattern in each scale and to exclude in Scale I one statement of a general nature. The scale analysis values for the resulting 10-item scales are for Scale I: H=.41; Delta-star=16.7; Rho=.78; and for Scale II: H=.38; Delta-star=17.8; Rho=.78. Thus Scale I can be supposed to measure the extent to which people worry about themselves, and Scale II to measure the extent to which people worry about others.

The two scales can be considered to be sufficiently comparable because it turns out that 6 of the 10 statements of each scale belong to related pairs. The fact that the two scales have no statements in common indicate that the tendency to worry about oneself and the tendency to worry about others reflect different attitudes.

For each respondent scores were established on the two scales. The responses on the statements were coded in such a way that the response category indicating a tendency to worry received one point, for the other response category no point was given. Thus the scores on each scale range from 0 to 10, the higher scores indicating a tendency to worry more strongly.

The respondent's scores on the two scales appeared to be positively correlated to a significant degree (Spearman rank correlation: $r=0.40$, $p<.01$). Hence we may conclude that people who tend to worry more strongly about themselves are also inclined to worry more strongly about others. Thus these data suggest that the popular belief which considers some people as real 'worriers' and others as 'carefree people' seems to be justified.

For each respondent comparable scores were obtained as regards the extent to which they worry about others and about themselves. Of the 100 subjects 68 received a higher score on Scale II (indicating that they worried more strongly about others), 9 subjects scored equally high on both scales, and 23 subjects received a higher score on Scale I (indicating that they worried more strongly about themselves). Of the 50 females with different scores on the two scales 41 (82%) appeared to worry more strongly about others than about themselves, whereas of the 41 males with different scores on the two scales 27 (66%) received higher scores on Scale II. However, this difference appears to be not significant (Chi-square=2.3, df=1). Based on these figures it can be calculated that if a sample is taken in a population with an equal number of males and females one would expect to find that of the people who worry more strongly about the others 55% would be female and 45% would be male.

Of the respondents below 30 years of age 67% worry more strongly about others than about themselves, while for the older respondents this percentage increases to 85%. However, here again the difference does not reach significance (Chi-square=2.7, df=1).

In order to compare the distribution of the scores over the two scales each scale was divided into three categories. Low strength of worry is arbitrarily defined as scale scores 0 to 3. Average strength of worry is defined as scale scores 4 to 6 while high strength of worry includes the scale scores 7 to 10. Table 5 presents the distribution of the scores for the two scales.

TABLE 5
Distribution of scores for scales I and II

Respondents with	low	average	high scores
Scale I	62	25	13
Scale II	33	34	33

The distributions appear to differ significantly (Chi-square=18.9, df=2, $p<.01$). The distribution for the scores of Scale I is strongly skewed indicating that most people do not worry strongly about themselves. The scores for Scale II are much more evenly distributed over the categories. This finding combined with the observation discussed above that the majority of the respondents score higher on Scale II than on Scale I indicates that our respondents tend to worry more strongly about others than about themselves.

When the relationship is studied between the nature of the worry and the scores on the two scales, no clear pattern emerges. From the data relating 'event' with the scores on Scale I (strength of worry about oneself) for only problems related to the respondents themselves, it appears that respondents with low scale scores (scores 0 to 3) mention most frequently problems related to daily matters, work and school, and problems of a psychological nature. However, respondents who worry strongly about themselves (scores 7 to 10) also mention these types of problems most frequently. When 'events' are related to the scores on Scale II (strength of worry about others) for only worries concerning others an equal lack of pattern is found. As could be expected the few respondents who worry about the possible death or serious illness or accidents of others tend to have high scale scores (but by no means the highest). Nevertheless, compared to the relatively carefree people the respondents who worry strongly about others mention relatively more frequently problems of a psychological nature. If in the two tables the rows with only zero frequencies are excluded the chi-square values appear to be insignificant, suggesting that the nature of the problem and the strength respondents worry about others are generally unrelated. (But for one table this analysis is dubious because of the relatively high number of low expected values).

From these data it seems safe to conclude that as regards the extent people worry, individual differences play an important role. Respondents tend to worry more strongly about others than about themselves. The nature of the problem has little effect on the strength with which people worry. Hence as regards the extent people worry the person involved in the worry seems to be of more relevance than the event involved.

SUMMARY OF RESULTS

From the two studies it appears that compared to the other categories worries more frequently concern daily matters. In addition, the respondents more often worry about him- or herself. In the second study it was even found that over half of the problems described concerned the respondent only. In both studies a significant relationship was observed to the effect that younger people tend to worry relatively more often about themselves, whereas older people tend to worry relatively most frequently about others, especially about relatives. The frequency with which respondents worry about a particular problem is most influenced by the nature of problem it is rather than by the person involved.

Based on the analyses concerned with the extent to which people worry about themselves and about others, it appears that in general respondents worry more strongly about others than about themselves. According to the results of the scale analysis the tendency to worry about oneself or to worry about others reflect different dimensions of the personality. Nevertheless a positive correlation is observed between the scores on the two scales indicating that some people tend to be rather carefree and that others tend to be true 'worriers'. In addition the extent to which one worries appears more dependent on the person involved than on the nature of the event. Strength of worry does not appear to be significantly related to gender or age.

Hence as a general conclusion one could say that people seem to be more frequently involved and worried about their own affairs, especially about daily matters, but that if they worry about others then they tend to take it more seriously and to worry more strongly about it.

THE TENDENCY TO WORRY AND SPES

The 'worry' hypothesis assumes that a direct relationship exists between the tendency to worry and the tendency to report SPES. In previous studies it has been shown that the characteristics of SPES have remained relatively stable over the past hundred years and over several Western cultures. Thus if the hypothesis is correct, the various properties of the tendency to worry can be expected to match more or less the known characteristics of SPES.

As observed in the introduction three main aspects can be distinguished about the tendency to worry: the event or situation one worries about, the person involved and the extent to which one worries about the problem. Hence for the evaluation of the 'worry' hypothesis the data of the two studies about these three aspects have to be compared with the data of the studies on SPES.

One discrepancy between worries and SPES which immediately appears concerns the category of events 'psychological problems and relational problems'. SPES nearly always involve specific events and rarely general impressions as for instance: "I have the paranormal impression that my partner and I are going to have trouble. Therefore, when the data on events as obtained in the present studies are compared to the data on SPES, the distributions are presented both including and excluding the category 'psychological problems'. For this comparison only the data of the 'Sannwald' and 'Rhine' collections are used, because the cases of the 'Phantasms' collection only involve SPES related to others whereas in the present studies the events concerned both worries about others and about the respondents. The distributions of events from collections of SPES and from the present studies are presented in table 6.

TABLE 6
Distribution of events

	death serious accidents %	light illness accidents %	material matters %	daily matters %	psychol. problems %	general events %	n
Rhine	53.3	18.1	3.3	21.2	--	4.1	1616
Sannwald	57.0	8.4	2.3	23.1	--	9.2	783
Study I	16.6	6.8	13.2	27.8	22.0	13.7	205
Study II	4.7	12.4	8.8	36.5	24.1	13.5	170
without the category 'psychological problems'							
Study I	21.2	8.8	16.9	35.6	--	17.5	160
Study II	6.2	16.3	11.6	48.1	--	17.8	129

From the table it appears that only in the category 'light illness or accidents' percentage values are observed which lie in the range of the values observed for that category for SPES. In all other instances the range of percentages from the studies on SPES clearly differ from the percentages obtained in the present studies.

The distributions over the different categories of persons involved in SPES and involved in worries people have are presented in table 7. From table 7 it appears that in no instance do the percentages for a

TABLE 7
Distribution of persons

	self %	family %	other relatives %	friends acquaintances %	general %	n
Rhine	20.0	49.6	9.5	16.7	4.2	1614
Sannwald	21.7	38.7	7.6	22.9	9.1	789
Study I	38.9	26.9	7.2	7.2	19.7	208
Study II	57.5	16.2	2.4	9.0	15.0	167

category of persons involved in worries lie within the range of the percentages of the same categories for SPES. The only agreement between the data appears to be that in the two sets of data a tendency can be noted that most experiences relate to other persons. A more detailed analysis of this finding will be presented when the effects of gender of the respondents or percipients is discussed.

In these studies the strongest relationship was observed between age and the tendency to worry. Unfortunately there is little known about the relationship between age and the occurrence of SPES, hence a comparison between the two phenomena as regards age effects is hardly possible. However, it is possible to study in more detail the effect of the sex of the person on the variables 'tendency to worry' and 'tendency to report a paranormal experience'.

When only cases are considered in which other persons are involved the percentage of female and male percipients in SPES are respectively 56% and 44% for the 'Phantasms' collection, 74% and 26% for the 'Sannwald' collection and 83% and 17% for the 'Rhine' collection. However, it is known that the cases for the 'Phantasms' collection where mainly acquired from a male-dominated environment whereas the data from the 'Rhine' collection were obtained from a female-dominated surrounding. These so-called 'reporting effects' have most probably increased the differences between the percentages of male and female percipients in the different collections.

The analogous data for the present studies are more in agreement with each other. Based on the number of problems mentioned by the respondents and assuming a population with an equal number of males and females from which a sample is taken of the people worrying about others, 57% would be female and 43% would be male according to Study I, and 59% would be female and 41% would be male according to Study II. The same analysis based on the scale scores reveal that, of the people worrying more strongly about others than about themselves, 55% would be found to be female and 45% would appear to be male. Although these percentages lie at the extreme of the range of values observed for the male-female distributions for SPES, these data can be considered as supportive for a possible relationship between the tendency to worry and SPES.

However, when the same analysis as regards sex differences is made for only events in which the person him- or herself is involved, an opposite result is obtained. From the persons who report a SPES related to themselves, 56% is found to be female and 44% to be male (Sannwald collection). The comparable figures for the Rhine collection are 67% females and 33% males. But in the present studies it was found, again assuming that a sample is drawn from a population in which males and females are equally distributed, that of the persons who report worrying about matters related to themselves, we would expect 33% to be female and 67% to be male (Study I) or 45% to be female and 55% to be male (Study II). The same analysis based on the scale scores yield expected values for people worrying more strongly about themselves than about others of 35% for females and 65% for males. Hence in these data an opposite trend is observed for the tendency to worry and for SPES. Thus the variable which yielded supportive evidence for the hypothesis in the previous section yield conflicting evidence when the same analysis is made for cases in which the respondents themselves are involved in the cases.

From the analyses presented above it can be concluded that in general the data on the tendency to worry lend little support to the 'worry' hypothesis for explaining SPES. The two types of experiences have in common that, in both, mainly events of a negative nature are involved. However, the distributions of events and persons involved in worries and SPES differ considerably. The only trend which appears common to the two phenomena is the finding that women tend to worry more often about others and also report more often SPES in which others are involved. On the other hand, the same variable yields conflicting evidence for cases in which only the respondent is involved. Whereas males tend to worry more frequently about themselves SPES related to the respondent are more frequently reported by females.

Because SPES as well as worries are most often of a negative nature it might be possible to salvage the hypothesis by introducing some restrictions. From the distributions of events (table 5) it appears that the greatest discrepancy between the data from this study and the data from SPES are observed in the category 'death and serious illnesses or accidents'. Thus a possible restriction could be that only worries about serious events result in the attribution of a paranormal character. In that case, however, it becomes difficult to explain why SPES hardly ever involve serious material events and why still a not insignificant number of these experiences involve trivial and positive events.

However, a proponent of the 'worry' hypothesis might attempt to salvage the hypothesis by further restricting it by assuming that only under specific circumstances, for instance, an emotional crisis in a person's life, feelings of worry are labelled 'paranormal' when the feared event becomes true. If this assumption holds one would expect that such circumstances increase the tendency 'to take refuge' in the paranormal. But so far the studies on factors influencing the belief in the paranormal do not provide us with data suggesting that specific circumstances would have such an effect on people. Thus salvaging the 'worry' hypothesis by introducing further restrictions seems, at least on the surface, not very promising. Therefore the data of these studies suggest that the 'worry' hypothesis is not likely to provide a satisfactory explanation for the phenomenon of SPES.

ABSTRACT

Spontaneous paranormal experiences (SPES) are experiences in real-life settings to which people attribute a paranormal character. Important differences between experimental research in parapsychology and the study of SPES are that SPES can be studied as a distinct class of human experiences which do not require the a priori assumption of the existence of psi and, further, that the study of such experiences allows one to falsify the concept of psi as an explanation for them.

It is possible to test the applicability of hypotheses of a non-psi nature which could lead to the conclusion that the concept of psi as an explanation for SPES has to be rejected.

The aim of the present studies was to find preliminary evidence for or against the validity of the 'worry' hypothesis. The 'worry' hypothesis associates SPES with the tendency of people to worry about possible future events. It assumes that a paranormal character is attributed to feelings of worry in the case that events people consider unlikely and about which they worried do come true.

A preliminary impression of the validity of this hypothesis can be obtained by comparing characteristics of the tendency to worry with the same characteristics for SPES. In two studies data were obtained from different samples and with different methods concerning the events and the persons people worry about. In addition scales were constructed to measure for each respondent in a comparable way the extent to which they worried about others and about themselves.

When these data were compared to the known distributions of the same variables for SPES it turned out that the differences are quite substantial. Hence the data of these studies suggest that the 'worry' hypothesis is not likely to provide a satisfactory explanation for the phenomenon of SPES.

REFERENCES

Kingma, J. 'A comparison of four methods of scaling for the acquisition of early number concepts', J. Gen. Psychol., 110, 1984, 23-45.

Mokken, R.J. 'A theory and procedure of scale analysis', Den Haag, Netherlands, Mouton, 1971.

Schouten, S.A. 'Analysing spontaneous cases, Research Letter Parapsychology Laboratory, University of Utrecht, 9. 1979, 55-63.

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

Schouten, S.A. 'Analysing spontaneous cases: A replication based on the Sannwald collection', E.J.P., 4, 1981, 9-49.

Schouten, S.A. 'Analysing spontaneous cases: A replication based on the Rhine collection', E.J.P., 4, 1982, 113-159.

Schouten, S.A. 'A different approach for analyzing spontaneous cases: With particular reference to the study of Louisa E Rhine's case collection', J.o.P., 47, 1983, 323-341.

Schouten, S.A. 'Quantitative Analysen paranormaler Spontanberichte' Z. Parapsychol. Grenzgeb. Psychol., 25, 1983, 1-27.

ON THE ASSESSMENT OF EVIDENCE FOR PSI

Nils Wiklund
Karolinska Institutet
Stockholm

Leading parapsychologists have made selections of important parapsychological experiments. The resulting list of experiments can be used for the assessment of the evidence for psi. This approach can yield both a "Yes" and a "No" to the question whether psi exists or not (in contrast to other approaches used by parapsychologists). It is not an easy task to assess the strength of the present-day experimental evidence for the existence of psi phenomena (i.e. telepathy, clairvoyance, precognition, and psychokinesis). Some scientists, notably parapsychologists, seem to consider the evidence to be fairly strong; others are not at all impressed. The primary question in parapsychological research is whether psi phenomena exist or not. If the answer is yes, this can be established either by a repeatable experiment or by individual experiments that are so well controlled that no other explanation is found. If the answer is no, neither method can be very efficient in procuring an answer. If psi does not exist a repeatable experiment that demonstrates psi cannot be found, and individual experiments containing no artifacts will necessarily yield chance results. Chance results, however, do not immediately prove the non-existence of psi, and so these two approaches will not yield an answer to the existential question in case psi does not exist.

The question of the existence of psi is sometimes neglected by parapsychologists. The majority of those who carry out experiments in parapsychology seem to work under the tacit assumption that paranormal phenomena exist (cf. Schmeidler, 1971; Beloff, 1973). And yet very few systematic studies have been made to assess the strength of the evidence for psi. More important, the conditions under which parapsychologists would be prepared to abandon the hypothesis that psi exists have not as yet been defined. This means that the assumption that psi exists has a certain similarity with an unfalsifiable assumption.

To take the existence of psi for granted is associated with various types of risks. One is that scientists in other fields will continue to ignore parapsychology (Pratt, 1979, p. 24) and that "We shall very quickly degenerate into an insignificant clique of tiresome pseudoscientists" (Beloff, 1973, p 190). Another closely related risk is that slack experimental procedures may be re-introduced into parapsychological research. One aspect of this is that parapsychologists may develop a certain blindness to alternative explanations to the unidentified experimental effects (interpreted as psi) often found in parapsychological research. From a methodological point of view it may actually be advantageous to use the assumption that psi does not exist, since this will keep the researcher most alert to the various artifacts that may be involved in a particular experiment. Perhaps a term that does not imply an interpretation would be helpful in this respect. ESP, both in the sense of 'Extra-Sensory Perception' and in the sense of 'Error Some Place' (Honorton, 1975) implies a rather specific interpretation. A descriptive term like 'Unidentified Experimental Effects' (UEE) would be more neutral.

One method to make an assessment of the evidence for psi is to select a number of important experiments and to expose these to a thorough scrutiny. This method has the potential of yielding both a "Yes" and a "No" to the existential question (in contrast to the two methods mentioned above). If the true answer is "No", it is likely that artifacts will be found. If the true answer is "Yes", artifacts will not necessarily be found.

One necessary prerequisite for this method is to decide which experiments are to be considered as important. It is quite easy to reveal faults in the reports of parapsychological experiments, but in the absence of a decision as to which experiments are important the only conclusion is that each faulted experiment does not carry any weight for the existence of psi, but then there are an endless number of other studies ... This is similar to the mystical description of God's nature in negative terms "neti, neti -not this, not this". Without a decision as to which experiments are important the dialogue with scientific colleagues in other fields would come to a standstill, but parapsychologists traditionally take a certain pride in welcoming such dialogues. It is now more than a century since the Society for Psychical Research was founded in 1882, and from this time awards parapsychologists have been able to keep this dialogue alive and to achieve a certain respectability. One essential aspect of this dialogue has been the presentation of lists of important experiments, e.g. by Rhine et al. (1940). During a period these experiments were even considered to be crucial for the assessment of the existence of psi. As long as no flaws had been revealed the weight of the evidence was rather heavy. Later, artifacts, or the possibility of artifacts, have been demonstrated in some of these experiments, and now the weight of the evidence, of course, appears much lighter. During the same time new important experiments have been conducted, together with many less important ones, and it is now necessary to make renewed decisions as to which experiments are at present to be considered as particularly evidential.

The purpose of the present study is to procure such a list of important experiments. I have asked a number of leading parapsychologists (former presidents of the Parapsychological Association) to make individual selections. Most of them have fulfilled my request, and this shows that parapsychologists' interest in communication with scientists in other fields is still as strong as it has been from the start. This list will be helpful for those wishing to make an assessment of evidence for psi.

PARTICIPANTS

Presidents of the Parapsychological Association during the period 1968 to 1978 were contacted for the study. In accordance with a suggestion from one of the ex presidents, J.G. Pratt was also contacted. It was reasoned that, even if Dr. Pratt did not happen to be president of the P.A during this period, he was without doubt one of the most important researchers in parapsychology. Since some of his experimental reports have been particularly scrutinised by those wishing to assess the strength of the evidence for psi (Hansel, 1980; Stevenson, 1967), Dr. Pratt's selection of important experiments may even be of special interest.

In a letter to the presidents I explained the purpose of the study and how they were supposed to make their selections. Here follows the central parts of the letter:

> "What I would like you to do is to specify which experiments (or cases) in your opinion constitute good, important studies, substantially contributing to the evidence for the existence of psi. The most important thing is that the experiments are methodologically sound. You must not be aware of any 'artifacts'. Otherwise I do not impose any limitation on your choices. You can mention any number of experiments, but your choices should be guided by the criterion that if artifacts are discovered in a particular experiment this would noticibly diminish the support for the existence of psi .... If you select a series of experiments because they seem to yield hope for a repeatable experiment please indicate in particular that this is the reason for your choice and also indicate which of these experiments in themselves furnish good evidence for psi irrespective of whether the experiment will turn out to be repeatable or not (again with the guiding principle that the evidence for psi would noticibly diminish if artifacts were discovered)."

The first letter to the former presidents of the P.A. was sent in 1979, and by 1981 I had received at least some kind of answer from ten of the twelve persons that I had contacted. A list of the participants is given below. The year of their presidentship is given in parentheses together with an abbreviation of their names.

Specific selections of parapsychological studies were made by the following: Ian Stevenson (I.S., 1968), Hans Bender (H.B., 1969), John Beloff (J.B., 1972), Robert L. Morris (R.L.M., 1974), Martin Johnson (M.J., 1976), K. Ramakrishna Rao (K.R.R., 1978) and J.G. Pratt (J.G.P.). (Note 1).

Others answered my request but did not make specific selections. Charles T. Tart (presidential period: 1977) wrote: "Generally speaking I would support the choices of other P.A. presidents as long as the reports were published in P.A. affiliated journals, but I would reserve the right to reject a specific experiment." Gertrude R. Schmeidler (presidential period: 1971) explained in detail her reasons for not making a specific selection of experiments; her opinion is discussed below. Robert L. Van de Castle (presidential period: 1970) wrote that he did not have the time to make the selection I wanted. Only two of the former P.A. presidents did not even return a card, on which I had asked them to state whether they intended to participate in the study or not (Rex G. Stanford and Charles Honorton - presidential periods: 1973 and 1975 respectively). It is improbable that none of my letters reached them, and I can only surmise that they consider their opinion about making specific selections of experiments to be sufficiently well known from the parapsychological literature (and that it is a negative one).

RESULTS

Eight studies (or groups of studies) were chosen by at least two of the P.A. presidents and they are listed here, starting with those studies that were mentioned most frequently. After each study is mentioned how many times it was chosen and by who. Twenty five studies were chosen by only one of the P.A. presidents and these will not be further discussed here.

(1) Helmut Schmidt's experiments using a binary random number generator.
(Five: J.B., R.L.M., M.J., K.R.R., J.G.P.)
(2) The Kanthamani & Kelly experiments with the subject Bill Delmore.
(Four: J.B., M.J., K.R.R., J.G.P.)
(3) The Stepanek experiments.
(Three: J.B., K.R.R., J.G.P.)
(4) The Ganzfeld studies.
(Three: J.B., M.J., R.L.M.)
(5) The Brugmans experiment at Groningen
(Two: I.S., J.B.)
(6) The Pearce-Pratt experiments.
(Two: I.S., K.R.R. - not J.G.P.!)
(7) The Pratt-Woodruff experiments.
(Two: I.S., J.G.P.)
(8) The Roll & Klein experiment with the subject Harribance.
(Two: J.B., K.R.R.)

Diana Robinson (1981) used a different approach in order to get a list of important experiments. She contacted all the members of the Parapsychological Association and asked them to make selections of parapsychological experiments. It is interesting to compare her results with the choices made by the P.A. presidents. It turns out that her final list of individual experiments or groups of experiments (Robinson, 1981) includes the first seven choices of the P.A. presidents, but not the last one. (I interpret her item 'Pratt's work in general' to cover the Pratt-Woodruff series.) In addition Robinson's list includes three other groups of experiments: the so-called remote viewing experiments (associated with the names of Targ, Puthoff, Bisaha, and Dunne), Schmeidler's sheep-goat work, and the Maimonides dream work. None of these was chosen by anyone of the P.A. presidents who participated in my study (although the Maimonides dream work was given an 'honorable mention' by one of them, I.S.).

Since the first seven choices of the P.A. presidents coincide with the choices made by parapsychologists in general, I think we can consider these seven groups of experiments to represent important parapychological studies.

REFERENCES

An introductory bibliography to the first seven groups of studies selected by the P.A. presidents is presented here. It is not meant to be complete, but the references will hopefully be a useful starting point for those who want to make an assessment of the studies.

(1) Helmut Schmidt's experiments.

Schmidt, H. 'Precognition of a quantum process', J.o.P, 1969, 33, 99-108.

Schmidt, H. 'Clairvoyance tests with a machine', J.o.P, 1969, 33, 300-306.

Schmidt, H. 'PK tests with a high-speed random number generator, J.o.P., 1973, 37, 105-119.

(2) The Kanthamani & Kelly experiments.

Kelly, E.F., & Kanthamani, B.K. 'A subject's efforts towards voluntary control', J.o.P., 1972, 36, 185-197.

Kanthamani, H. (B.K.), & Kelly, E.F. 'Card experiments with a special subject. I. Simple clairvoyance', J.o.P., 1974 , 38, 16-26.

Kanthamani, H. (B.K.), & Kelly, E.F. 'Awareness of success in an exceptional subject', J.o.P., 1974, 38, 355-383.

Kelly, E.F., Kanthamani, H., Child, I.L., & Young, F.W. 'On the relation between visual and ESP confusion structures in an exceptional subject', J. A.S.P.R., 1975, 69, 1-32.

(3) The Stepanek experiments.

Pratt, J.G., Stevenson, I., Roll, W.G., Blom, J.G., Meinsma, G.L., Keil, H.H.J. & Jacobson, N. 'Identification of concealed randomized objects through acquired response habits of stimulus and word association', Nature, 1968, 220, No. 5162, 89-91.

Blom, J.G. & Pratt, J.G. 'A second confirmatory ESP experiment with Pavel Stepanek as a "borrowed" subject', J. A.S.P.R., 1968, 62, 28-45. See also letter by Pratt in J. A.S.P.R., 1969, 63, 207-209.

Pratt, J.G. 'A decade of research with a selected ESP subject: an overview and reappraisal of the work with Pavel Stepanek', Proceedings of the A.S.P.R., 1973, 30, 1-78.

Keil, H.H.J. 'Pavel Stepanek and the focussing effect', Research Letter of the Parapsychology Laboratory, University of Utrecht, No. 8, Oct 1977.

(4) The Ganzfeld studies.

Honorton, C. 'Psi and internal attention states', in Wolman, B.B (Ed.), Handbook of Parapsychology, 1977, Part V, Chap. 1.

(5) The Brugmans experiments at Groningen.

Brugmans, H.J.F.W. 'Une communication sur des experiences telephatiques au laboratoire de psychologie a Groningen faites par M. Heymans, Dr. Weinberg et Dr. H.I.F.W. Brugmans', Le compte rendu officiel du premier congres international de recherches psychiques, Copenhagen, 1922, 396-408.

Murphy, G. 'Challenge of Psychical Research', New York, Harper, 1961, pp. 56-62.

Pope, D.H. 'The Brugmans experiments', J.o.P., 1952, 16, 1-3.

Schouten, S.A., & Kelly, E.F. 'On the experiments of Brugmans, Heymans and Weinberg', E.J.P., 1978, 2, 247-290.

(6) The Pearce-Pratt experiments.

Rhine, J.B. & Pratt, J.G. 'A review of the Pearce-Pratt distance series of ESP tests', J.o.P., 1954, 18, 165-177.

Hansel, C.E.M. 'ESP and parapsychology. A critical re-evaluation', Buffalo, N.Y., Prometheus Books, 1980, Chap. 10.

(7) The Pratt-Woodruff experiments.

Pratt, J.G. & Woodruff, J.L. 'Size of stimulus symbols in extrasensory perception', J.o.P, 1939, 3, 121-158.

Pratt, J.G. 'New evidence supporting the ESP interpretation of the Pratt-Woodruff experiment, J.o.P, 1976, 40, 217-227.

Hansel, C.E.M. 'ESP and parapsychology. A critical re-evaluation', Buffalo, N.Y., Prometheus Books, 1980, Chap. 11.

COMMENTS

It is of particular interest that the Pearce-Pratt and the Pratt-Woodruff experiments are still considered to be good evidence for the existence of psi by many parapsychologists. These experiments have been carefully examined by various investigators, and the assessment of these studies is therefore, greatly facilitated. It is also of principle importance to see if any parapsychological experiment can stand the test of a thorough critical examination. We have here then two series of experiments that some parapsychologists consider to have passed the test. J.G. Pratt, who was involved in both, mentioned only one of them though (the Pratt-Woodruff experiments). Helmut Schmidt's experiments have also been examined by Hansel in the second edition of his book (1980). However most of the P.A. presidents made their choice of Schmidt's experiments before this edition of the book appeared.

The Ganzfeld studies seem to have been chosen not primarily because they represent particularly well designed experiments from a methodological point of view but because they "have a better record for repeatability than any other type of experiment in recent years" (i.e. in parapsychology) (Beloff, 1980, p. 118).

The importance of a list of experiments like the present one becomes evident when it is compared with the list of studies examined in the third part of Hansel's book (1980) where he attempts to cover the developments in parapsychological research since 1965. He considers "the use of machines and automated procedures" to be the most important development since 1965, and therefore he devotes a full chapter to the Helmut Schmidt experiments, which are considered to be important also by the P.A. presidents. Apart from this, Hansel's choices of experiments do not coincide very much with those chosen by the ex-presidents. He devotes a full chapter to psychical research at SRI, telepathy and dreams, Uri Geller, the remote viewing experiments, and what he calls "the challenge of chance". None of these studies was selected by even one of the P.A. presidents as important studies. In one chapter on "The miracle men" he does, however, devote three pages to the Stepanek experiments. The Kanthamani & Kelly experiments as well as the Ganzfeld studies are not at all examined. Honorton (1981) is undoubtedly correct in his criticism that "Hansel's coverage of the recent research is highly selective and grossly inadequate" (p. 155). But how could Hansel know which recent studies to examine when parapsychologists nowadays according to my view often are less eager to specify important experiments than they were in earlier decades? When Hansel's choices are compared with those of parapsychologists in general (Robinson, 1981) it is also found that two further chapters in Hansel's book are devoted to what they consider to be important studies, although these studies are not selected by the P.A. presidents (the remote viewing experiments and the Maimonides dream work). And one cannot deny that the rest of Hansel's subject areas have received wide publicity in recent years.

DISCUSSION

There are basically two types of objections to the idea of making a list of particularly important parapsychological experiments. The first objection is that evidence should not be derived from individual experiments but from the combination of all experiments, particularly as they support each other in a systematic (or theoretic) way. This objection is commonly known as the 'fagot' theory. The second objection is that individual experiments cannot, in principle, yield conclusive evidence; only a repeatable experiment can do this.

The first objection was explained in a sympathetic way by Gertrude Schmeidler (1979) in a letter to me. She wrote:

> "You wanted a list of experiments in parapsychology, or of cases, on which (here I rephrase what you wrote) conviction depends. But to my mind, conviction grows slowly and becomes strong and complete only when many, many data accumulate to form a pattern ... conviction comes when a new theory or pattern begins to emerge from diverse findings, and when predictions from a theory are supported by the outcome of research - especially of independent research. It adds up to a network of interrelated findings which begin to form a meaningful pattern. And this means that a (linear) list of separate experiments, each of which seems good, isn't an adequate answer to your basic question ... Any one experimental series might contain artifact or be due to fraud. Our own ingenuity may be insufficient to spot the artifact. But if data obtained by varied procedures support each other, the artifact of one experiment would not appear in others (which change the particular procedures). If we spotted an artifact, we've mentally discarded that experiment; if we doubt our ability to spot all responsible artifacts, we're reassured by systematic rather than direct replication."

However, to insist on the cumulative evidential value of all parapsychological experiments will hardly convince the sceptic. Price

(1955) firmly rejected the idea that the 'mass' of experiments carries any evidential value.

> "But the answer that will impress me is an adequate experiment. Not 1000 experiments with 10 million trials and by 100 separate investigators giving total odds against change (sic!) of 10 to 1 - but just one good experiment. And until such a demonstration has been provided, I hope that my fellow-scientists will similarly withhold belief."

Some of the parapsychologists who insist on the cumulative evidence of all experiments seem to suggest that one should not be preoccupied with isolated studies, because, so far, no individual experiment has been found to be without errors. They suggest that there probably is no impeccable experiment and, even, that there cannot be one. This opinion, however, rather amounts to saying that there is no strong evidence for psi, and that there cannot be, which comes very close to saying that psi does not exist.

The second objection to the idea of making a list of important parapsychological experiments is that individual experiments cannot in principle prove the existence of psi. Only a repeatable experiment has this capacity. This approach, however, is of little help in assessing the evidence for psi in the absence of a repeatable experiment. Many leading parapsychologists seem to be of the opinion that there is at present no genuinely repeatable experiment. Nevertheless they do not draw the conclusion that paranormal phenomena do not exist. The absence of a repeatable experiment is not interpreted as a "No" to the question whether psi exists. The search for a repeatable experiment is, therefore, of limited value when trying to make an assessment of the evidence for psi. This approach can yield an answer to the question whether psi exists only under the condition that psi does really exist (and a genuinely repeatable experiment would, of course, make the existence of psi uncontroversial). (Penetrating studies have also been made of the very concept of repeatability, e.g. by Bauer, 1979, but a discussion of these studies lies outside the purpose of the present paper.)

Many parapsychologists seem to work under the assumption that psi phenomena exist, and some even seem to take their existence for granted. If parapsychologists in general have not seriously considered the possibility of the non-existence of psi, I think this would introduce a considerable methodological danger. There are, in fact, signs of an apparent methodological crisis in parapsychology. Those who take the existence of psi for granted tend to become less aware of the 'anomalous' and 'revolutionary' character of psi phenomena, in Kuhn's (1962) sense of the terms. As a consequence of this there is the risk of increased blindness for artifacts in the experiments. There is also a tendency to diminish the rigour of the experimental methods. In the 1950's the so-called two-experimenter design was proposed as a necessary part of a well conducted parapsychological experiment. Nowadays it is rarely used. Some parapsychologists stress the notion that no individual experiment can be 'perfect', and take this as an excuse to re-introduce slack experimental procedures. Nomina sunt odiosa. These developments represent a real methodological crisis in parapsychology and warnings have been voiced lest parapsychology will develop into a pseudo-science and "pass more and more into the hands of cranks as its present relatively respectable supporters drop away" (Alan Gauld, quoted by Nicol, 1980). It will not be easy to produce infallible and complete antidotes to this crisis. Much can be learned from J.B. Rhine's attitude to criticism from his scientific colleagues:

> "Critics both within and without parapsychology examined the first work and found, or thought they found, methodological loopholes in it. Rhine welcomed all such criticism and transformed them into constructive contributions: with the help of his colleagues he invented or utilized new methods which would make the research fully rigorous." (Schmeidler, 1981, p. 19)

Not only is it necessary to welcome critical examinations from colleagues in other fields, it would also be favourable if parapsychologists took a more determined responsibility for conducting the critical analyses of 'claims of the paranormal', and for disseminating the results of these analyses. It is an unfortunate development if this task is delegated to groups of researchers independent of the Parapsychological Association. I am favourably impressed by some leading parapsychologists for their scientific ability and for their immediate perception of weaknesses in experimental reports, for example at the P.A. conventions. But it is hardly enough to confine these critical analyses to oral discussions at these conventions. There is a risk that obviously deficient experiments will be considered to be adequate by parapsychologists in general. This may be an explanation of some of the differences between the selections of important experiments made by leading parapsychologists as compared to the choices by parapsychologists in general (e.g. concerning the remote viewing experiments). Bauer (1979) stressed that parapsychologists traditionally are the most competent critics of parapsychological experiments, and it seems vitally important to keep this tradition alive.

There is also another development in parapsychology that has been called a crisis. It is concerned with the weight of the evidence for psi. Scriven (1961) described the beginning of this 'crisis' which is characterized by a continuous "erosion of the evidence" for psi. (Cf. Bauer, 1979) This erosion of the evidence has continued since then, and the weight of the evidence for psi now appears less than it was before. But this development only shows the strength of parapsychology as a science. It shows that the scientific methods of parapsychology are still functioning healthily. This so-called crisis, concerned with the diminished weight of the evidence for psi, is therefore no real crisis, but merely a new phase in the development of parapsychology as a science.

As researchers in this field we are not necessarily supposed to prove the existence of psi. It is not even more desirable that paranormal phenomena exist than is their non-existence. Even less is there any merit in belief or disbelief in the phenomena (Beloff, 1973). We are merely here to investigate whether the phenomena exist or not. Depending on the answer to this basic question we can then attempt to find explanations. If the answer is yes (or judged to be yes) it is natural to search for a theory of psi. If the answer is no (or judged to be no) it will be necessary to find explanations to the unidentified experimental effects (UEE) in numerous parapsychological experiments, which will probably be of interest from a methodological point of view. It will also be necessary, and perhaps more interesting from a psychological point of view, to find explanations for the frequently reported spontaneous phenomena of a seemingly paranormal nature, that often carry a strong flavour of conviction for the person experiencing the phenomenon. If these experiences are not psi occurrences, then what are they?

NOTES

1. All of these, except Dr. Bender, completed their selections already in 1979.

Due to an unfortunate incident Dr. Beloff's list was seen by Dr. Rao and by Dr. Pratt before they completed their lists in the fall of 1979. Dr. Rao has informed me that he does not think he was influenced by Beloff's choices, and supported this view with references to his earlier publications. This issue is not essential however, since the consensus I was looking for anyhow could be expected to depend on marked interdependence between the opinions of the participants, due to their intimate knowledge of the research field. Later I also contacted the presidents for 1980 and 1981, but both decided not to make a selection of the experiments. Dr. Child thought that he would be too much influenced by Dr. Beloff's list that he had already seen, and Dr. Palmer argued that the entire parapsychological literature must be considered.

REFERENCES

Bauer, E. 'Kritik und Kontroversen der Parapsychologie', In: Die Psychologie des 20. Jahrhunderts, XV, p 546-559. Zurich, Kindler Verlag, 1979.

Beloff. J. 'Belief and doubt', Research in Parapsychology 1972, Metuchen, N.J., Scarecrow Press, 1973, p. 189-200.

Beloff, J. 'Seven evidential experiments', Zetetic Scholar, July 1980, 91-94, 116-120.

Crumbaugh, J.C. 'A scientific critique of parapsychology', In Schmeidler, G (Ed.): 'Extrasensory perception', New York, Atherton Press, 1969, p. 58-72. (Reprinted from: International Journal of Neuropsychiatry, 2, 1966, 523-531.)

Hansel, C.E.M. 'ESP and parapsychology. A critical re-evaluation', Buffalo, N.Y., Prometheus Books, 1980.

Honorton, C '"Error some place!"', Jounal of Communication, 25, 1975, 103-116.

Honorton, C 'Has science developed the competence to confront claims of the paranormal?', Research in Parapsychology 1975, Metuchen, N.J., Scarecrow Press, 1976, p. 199-223.

Honorton, C. 'Beyond the reach of sense: Some comments on C.E.M. Hansel's ESP and parapsychology: A critical re-evaluation', J. A.S.P.R., 75, 1981, 155-166.

Johnson, M, 'Interview with Martin Johnson in 1979', (Partly published in Sokaren, 1981, in Swedish)

Kuhn, T.S. 'The structure of scientific revolutions', Chicago, University of Chicago Press, 1962.

Nicol, J.F. 'Comments to Beloff' (1980).

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

Pratt, J.G. 'Parapsychology, normal science and paradigm change', J. A.S.P.R., 73, 1979, 17-28.

Price, G.R., 'Science and the supernatural', Science, 122, 26 Aug. 1955, 359-367.

Rhine, J.B. et al. 'Extrasensory perception after sixty years', New York, Holt, 1940.

Robinson, D. 'Letter', June 17, 1981.

Schmeidler, G.R. 'Parapsychologists' opinions about parapsychology', 1971, J.o.P., 35, 1971, 208-218.

Schmeidler, G.R. 'Letter', April 15, 1979'

Schmeidler, G.R. 'Rhine's contributions to experimental methodology and standards of research', J.o.P., 45, 1981, 11-23.

Scriven, M. 'New frontiers of the brain', J.o.P., 25, 1961, 305-318.

Stanford, R. 'Comments to Beloff (1980).

Stevenson, I. 'An antagonist's view of parapsychology. A review of professor Hansel's ESP: A scientific evaluation', J. A.S.P.R., 61, 1967, 255-267.

Stevenson, I. 'Letter', May 7, 1979.

Nils Wiklund, Ph.D.
Radmansgatan 77
S-113 60 Stockholm
Sweden

EUR.J.PARAPSYCHOL., V.5, 1984

WHAT IS WRONG WITH THE DEFINITION OF PSI?

Walter von Lucadou

WELL-KNOWN PROBLEMS OF PARAPSYCHOLOGICAL DEFINITIONS

In the glossary of Wolman's 'Handbook of Parapsychology' (1977) psi is defined as: "a general term to identify a person's extrasensorimotor communication with the environment. Psi includes ESP and PK." ESP is defined in the following way: "Knowledge of or response to an external event or influence not apprehended through known sensory channels", and PK is defined as: "the influence of mind on external objects or processes without the mediation of known physical energies or forces."

It is well known that these definitions do not say what psi, ESP, or PK really is but merely what it is not. The negative character of the standard definition is widely criticized by many authors (see Mabbet, 1982, Bauer and Lucadou, 1979). It prevents parapsychology from being a normal science where positive definitions are usual. The 'problem of the negative definition' includes further problems: If, for instance, it turns out that an ostensible paranormal effect can be explained by our present knowledge in the natural sciences does the definition require that it is no longer a paranormal effect? The corpus of knowledge in normal science, however, is never closed so we get the 'problem of preliminary knowledge'.

It is obvious that we must exclude trickery or manipulation if we investigate psi phenomena, but what else should we exclude? For instance, should we shield the subject electromagnetically? We call this the 'problem of exclusion'.

Similar to the problem of exclusion, but more specific is the 'problem of the boundary conditions'. Many psi researchers believe they know that psi events mainly occur under certain conditions, for instance with puberty problems of a RSPK agent. There exists however many cases where the same conditions are given without any outbreak of the paranormal.

The last two categories of problems have a more general and epistemological character and are not primarily linked to the problem of the negative definition. The first is called 'the problem of coherence of the description language'. Assuming that psi effects really exist as parapsychologists understand them, we have to consider a physical effect, for instance a physical process (the distortion of a metal or the transmission of a piece of information) which is linked with psychological conditions or more generally with a person's psyche. To describe physical processes we normally use the description language called 'physics'. For psychological things we use the description language called 'psychology'. Many scientists believe that both languages could be unified at least in future. At the moment, however, we are far from the goal of this programme called 'reductionism' (see Lucadou and Kornwachs, 1983). Even if we could describe both the psychological and the physical part of a psi effect readily with their proper languages it is extremely difficult to specify the interface between the two domains. If, for instance, we observe that a psychic can bend a certain metal but cannot bend another material, we do not know whether this is due to the differences in material (which is a physical variable) or due to psychological conditions (for instance he or she likes gold but not iron).

Similar to the former problem, the second more general one already contains a central part of the next paragraph. It is called 'the problem of consistency'. It simply states that we do not know which psi phenomena belong together, so we do not know whether ESP and PK are caused by a single psi process or whether a poltergeist phenomenon is a special form of PK, called Recurrent Spontaneous Psychokinesis (RSPK). This problem arises from the necessity to use models in science. And even if all the other problems could be solved we would still have to struggle with the latter, but this problems exists in normal science too.

THE PROBLEM OF IMPLICIT MODELS

The problems with definition of psi mentioned in the last paragraph are more or less well known and in spite of the fact that many critics of parapsychology such as C. Hansel, M. Bunge or J. Alcock use these deficiencies to prove that parapsychology is a pseudoscience, most parapsychologists are well aware of the problems, but they have a quite different definition in mind when they speak of psi. But what do they have in mind? Probably it would fit with the idea of most parapsychologists to say that psi has something to do with the mind-body problem or with a yet undiscovered force or energy such as 'entelechy' introduced by H. Driesch. One could also say that psi is a phenomenon linked with psychophysical systems. These concepts are primarily thought to be positive counterparts of the negative definition, but as long as these concepts are not specified further they are not helpful. An unknown energy or force exhibits the same problems discussed above. Moreover such terms as 'entelechy' and others proposed have had no application in normal science until now, thus they would not eliminate the isolation of parapsychology. One should note, however, that normal psychology shared similar problems when dealing with concepts like 'psyche'. But today such terms are no longer necessary and are not found in modern textbooks.

Psychology has invented a method to circumvent such problems by the use of so-called 'operational definitions'. An operational definition simply describes the experimental situation. For instance IQ is defined by the questionnaires which are used to measure it. Of course this method was applied in parapsychology too. For instance, PK could be defined operationally by the description of a Schmidt-type PK machine with the subject sitting in front of it. The description of course must also contain the instruction given to the subject. Here, a new problem will arise. A skilled experimenter would not say to the subject: "Please try to do PK!" because he would realize that in this case the definition becomes circular and worthless. (But we will come back to this problem.) Normally the experimenter gives an instruction like this: Could you try to influence the machine in such a way that the lights will flip in the given direction! But without touching it, please!" At first it may seem that the problems of definition have now vanished. But what does the instruction "influence" mean? All the problems mentioned above are hidden in this single word. To put it more precisely: the problems of definition are now enclosed within the model adopted by the experimenter and enter the instruction given to the subject. We will call these hidden assumptions 'implicit models'.

These implicit models contain more than the simple assumption that there is an 'influence' of the subject on the physical process. This can be illustrated by the Rhinean experimental paradigm:

(a) Rhine considered psi as being an 'influence' which means an energy or force or information which can be transferred from the subject to another subject or to a physical process. ("Mind has a real force.") An 'entity' is 'going out'.

(b) Psi can be accumulated by statistical means and/or repeated observation. Otherwise it would not make sense to perform statistical experiments. But also with spontaneous cases one believes that 'evidence' can be accumulated by repeated observation.

(c) Especially with experiments it is suggested that the outcome of the experiment should not depend in principle on when and where the experiment is performed. Only the physical and psychological conditions of the actual situation should play a role and it is further assumed that these conditions can be repeated at least to some extent. Thus psi is considered as being independent of the location and time (date) if all other conditions remain constant. We call this ergodicity.

(d) It is assumed that psi can be separated in principle from its content or meaning. Otherwise it would not make sense to try to investigate ESP by card guessing experiments using standardized symbols. This does not mean, however, that the meaning or the content would not play any role, but they are considered as separable variables which determine the psi process. Thus emotional content is regarded as a positive amplifier to the process.

These implicit models seem to be so simple, so plausible, and evident that they are generally adopted within parapsychology and often they are regarded as the basic categories which are necessary to guarantee a scientific investigation of psi. Nevertheless they could be wrong, totally or partially. On the other hand such wrong implicit models could prevent us from discovering the real structure of psi. The widely known elusive and chaotic character of psi seems to indicate that something fundamental to our hitherto methods and assumptions might be wrong, but do we have alternatives?

WHAT IS WRONG WITH THE IMPLICIT MODELS?

If we should find some of our implicit assumptions wrong, this would not necessarily mean that we must remove them totally. Very often such a wrong assumption can still be used as a simplified approximation which still has its worth for a special purpose. The implicit use of the notion of absolute space and time in classical mechanics may serve as an example. The invention of the theory of relativity did not devalidate classical mechanics at all, but illustrated its conceptual and practical limitations. Implicit models become dangerous only if one is not aware of using them. Nevertheless one could also imagine that certain models are incompatible with others. Therefore one should be very careful in deciding that a certain fundamental notion must be given up. Very often laymen but also parapsychologists tend to throw the baby out with the bathwater if they discover such fundamental discrepancies. They often claim that psi is non-physical, or cannot be investigated by scientific means; new paradigms are required. Here we will pick up first a single assumption which seems less fundamental than the others, but we will find that when it changes the others will also change slightly. This indicates that these assumptions are not totally independent of each other.

First of all it seems plausible that the assumption (c) must be dropped since everybody knows that it is practically impossible to guarantee that an experiment does not depend on its location in space and time. This is part of the replicability problem in parapsychology which, however, I will not discuss here (see Hövelmann, this issue). It is by no means possible even to get an identical replication of an experiment. Supposing we could arrange the same experimental conditions as in an earlier experiment with the same subjects involved we could not neglect the history tied to the subject. The experiences he or she has had in the first experiment might have an important influence on the present one. Nevertheless we will not attack this assumption. The reason is twofold: First, the problem is well known and normally it will not be overlooked so that an estimation of its influence could be given under favorable conditions; second, we do not have an alternative possibility.

Obviously assumption (a) is of central importance. It contains the notion 'influence' which was even explicitly used in both the operational and the standard definition. This concept plays a predominant role not only in physics but also in all other natural sciences too and is of special relevance in cybernetics and general system theory. In physics the term 'influence' is tied to the term 'interaction' which means that a force is applied to something (for instance gravity to an apple). It can be represented by a field or by an exchange (flow) of (virtual) particles. It always causes a change in energy. The use of the term 'influence' in other natural sciences is very similar to its use in physics, but very often the related physical or chemical processes are much more complicated so that 'influence' is very often a rather global description which is used in the sense that a given variable 'influences' another one.

It was especially the merit of general system theory to deliver a more detailed description of the concept of influence. The first step was a generalization of the concept in the sense that it was not only tied to special physical interactions, but to the concept of information. This can also be illustrated with the definition of ESP. The following statements are equivalent: A person influences another person or; a person transmits information to another person. The concept of information entails the concepts of transmitter, receiver, and channel. They can be realized by quite different physical objects. A measure of information is given by the ratio of the expectation values of certain states (signs) of transmitter, receiver, and channel. But the term information is only an element of the term influence, it is the carrier of influence. Influence basically means 'control'. For instance if a subject wants to influence a random generator he or she wants to control it (to force it into a certain direction). Without a given goal an influence could not be detected. Sometimes the goal is not specified explicitly, but there must always be a criterion to measure the influence and this can be considered as the given goal. In general system theory the concept of control is widely developed due to its practical relevance (Kliemann and Muller, 1976). It is tied to three other notions which simultaneously decribe the system in question. They are called observability, reachability, and stability. For stationary systems one can express these terms by simple mathematical formulas. Such a system can be described by its input, output, and its internal states. Let $u(k) \in \mathbb{R}^m$ be the input vector, $x(k) \in \mathbb{R}^n$ the state vector and finally $y(k) \in \mathbb{R}^r$ the output vector of the system at a certain time k then the system is completely described by the following equation:

$$\begin{aligned} x(k+1) &= A\,x(k) + B\,u(k) \\ y(k) &= C\,x(k) \end{aligned} \tag{1}$$

A, B, and C are matrices which represent the system. With these equations one can calculate the state x and the behaviour y of the system at a time k+1 if all other values are known. Indeed, A, B, and C might have a very complicated structure and in many cases one does not know them. In real systems they might depend explicitly on time and then the system will no longer be stationary. But we will not discuss these problems here. Nevertheless we can use this example to elucidate our concept. Controlability means that one can influence the state x of the system and thus its output by manipulating the input vector u. Now different problems will arise. Even if I knew the matrices A, B, and C, we must first find out the present state of the system. This, however, can only be done by the output of the system. Now it depends on C in equation (1) whether we can infer x(k) by y(k). Those systems where this can be done are called 'observable'. Assuming this might be possible, we still do not know whether we will ever reach our goal by manipulating u(k). To visualize the problem it is usual to give a graphical representation of the system which is called phase space. It represents the state x of the system depending on its essential (canonical) variables k.

Figure 1 shows the trajectory of a pendulum in its phase space.

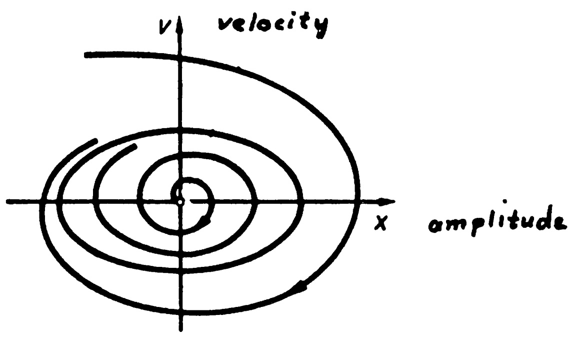

Figure 1

The trajectory of a pendulum in its phase space.

It may well be that a certain domain in the phase space cannot be reached by the system, for instance if the pendulum cannot surmount a certain amplitude or velocity. Reachability means that it must be possible to find a continuous trajectory or path from the starting point in the phase space to the goal. Finally stability means that the trajectories must not leave any upper boundary in the phase space.

These considerations show that the term 'influence' or 'control' can be specified and that there exist many possibilities where the use of these concepts can fail. It would be premature, however, to conclude that the known reliability and elusiveness of psi must necessarily indicate that these concepts can no longer be applied. We must try to test this experimentally.

But before we enter this part, let us take a brief look at the problem of learning. It is obvious that control is an essential feature within every process of learning. One needs both a goal which should be reached and a kind of feedback which indicates that the goal is approximated. If it should turn out that psi is not an influence, it follows that it is highly improbable that it could be enhanced by learning in the usual sense (see last paragraph).

The following PK experiment was performed to test among other hypotheses how a possible PK effect can influence a Schmidt-type PK machine. Obviously many of the published statistical PK experiments show a significant deviation of the scoring rate when a subject tried to influence it. Many experiments also show significant correlations to psychological variables such as the sheep-goat variable. The random event generators (Schmidt, 1974), which are normally used produce a sequence of equally distributed events by stopping a fast running flip-flop at a random moment, for instance when an impulse from a radioactive source is detected. With these random generators, however, it is impossible to figure out at which point of the machine the influence enters or which physical processes will be affected; it might be possible that the decay rate of the radioactive source could be changed; it might be possible that the sensitivity of the detectors could change; the clock that triggers the flip-flop could be influenced and finally the flip-flop or other electronic devices could change their behaviour.

Our experiment explicitly tested the question whether the radio-active source was influenced. In order to achieve this it was possible to construct a new type of random generator which has many advantages. A calculation shows that even with a highly significant result of a statistical PK experiment (e.g. CR=5) it is not possible to measure this effect directly as a net effect of delayed or shortened life time of the radioactive source. Thus we decided to install a radioactive source in the center of a circle drawn by five Geiger-Mueller counters. Only the random events of one of the counters was displayed to the subject, whereas the others were stored simultaneously. Assuming that the subject could influence the decay-rate of the counter one would expect that the other counters would also detect these perturbations. The technical details are described elsewhere (Lucadou, 1982). All experiments were run in a double blind situation so that no one actually knew which type of random generator was active. The whole experiment consists of two parts: the pilot study and the final experiment. During the pilot study the psychological and situative variables were described only qualitatively and the subjects could choose the length of the experiment. In the final experiment everything was fixed and the psychological variables were measured with standard questionnaires. The results of the final experiment will be described elsewhere. However, from the point of view of physics both parts of the experiment were identical. The results regarding the 'influence hypothesis' are also identical for both experiments. It turned out that there was no hint at all that the radioactive source itself was influenced. During the pilot experiment we very often obtained good results with single subjects scoring better than CR=3 and the situative and psychological conditions allow the experiment to be considered as a successful PK experiment. But there was no influence on the source. The technical details will be described elsewhere.

Obviously from this result alone one can not conclude that the concept of 'influence' could not be applied since one could imagine an ostensible PK effect on any of the other processes discussed above. To investigate this hypothesis we used a statistical technique which however does not allow equally precise answers as the technique described above. The usual evaluation procedure of parapsychology tests whether the gained scoring rate differs significantly from the null hypothesis, which means a chance result. But even if the subject should influence the scoring rate according to the instruction, it could well be that such an influence could be washed out by fluctuations so that the scoring rate would not be a good measure for such an influence. Thus it is necessary to combine several statistical procedures to find out what happened. I do not want to go too much into the technical details here; however, the method we used is rather new in our field so I must explain it at least very roughly. Instead of normal random sequences we used a Markoff-chain of '0' and '1' events. The parameters of this Markoff-chain are given by a transition matrix T:

$$T = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix} \qquad (2)$$

where p_{00} gives the probability to get a '0' and p_{01} the probability to get a '1' after a '0'. The other values are vice versa. The values of the matrix are exact and do not depend on any instrumental parameters or the distribution of the radioactive pulses. The only assumption which must be made is that the primary random process (the radioactive source) is really random, which means that the single events do not depend on each other. The autocorrelation function of this Markoff-chain only depends on the values of the matrix. Thus it is a good method to test the basic assumption and it turns out that it is very sensitive to any deviation from chance. But we also used several other statistical tests, which test deviations both from the theoretical distribution and from randomness. Then we compared all these tests with each other.

The results were surprising: even with those runs which showed a highly significant deviation in the scoring rate, the other procedures did not indicate any non-random influence on the sequence. All deviations which occurred looked like normal fluctuations which must be expected within stochastic sequences. However, they occurred more frequently than one would normally expect. It is necessary to elucidate this point a little bit further. Let us assume for example that the distribution function of a random variable is a gaussian. Then we know that the variance and the mean value of the sequence in question are independent variables. For other distributions this need not hold. Let us assume further that our original sequence will now be influenced by an arbitrary non-random process which changes the distribution of the sequence, for instance its mean value. Even if we do not know the actual disturbing process it is rather unlikely that the resulting deviation is again a gaussian. (Note however that R. Jahn also obtained a gaussian with the 'influenced' sequences which is in agreement with our results. (Jahn, 1982)) This is only the case if the 'influence' is random itself. But if we remember what we had learned about the notion 'influence', we should not call such a process 'influence'. Thus we would expect that the mean value and variance are no longer independent. A similar consideration holds for other statistical quantities. However if one does not know anything about the influencing process it is rather difficult to give a quantitative measure for its randomness. Therefore we simulated periodic, non-periodic, feedback driven, and non-feedback driven influences with a certain degree of randomness in order to find out how our criteria will react to it. Figure 2 shows the relationship between the scoring rate and the mean square deviation from the autocorrelation function of our Markoff-chains for all runs of the pilot experiment.

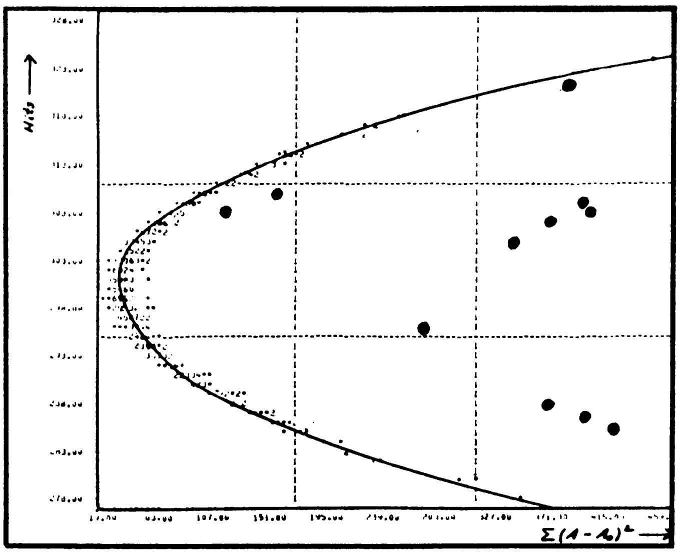

Figure 2
Relationship between scoring rate and mean-square deviation from the autocorrelation function of the Markoff-chains for all runs of the pilot experiment.

We see that the relationship has a quadratic shape. The dark dots stem from runs with simulated influence. They do not lie on the curve. All the other runs hit the curve more or less precisely, especially those with a high scoring rate which can be considered as successful PK runs. Therefore we are led to conclude that PK cannot be considered as an influence in the sense discussed above. One should note, however, that these results are preliminary since until now we cannot give a quantitative measure or a level of significance for this result, but the work is in progress.

Now we are confronted with a rather intriguing problem: on the one hand we have performed some successful PK experiments with statistically significant scoring rates (the best subject produced CR=3.2) which also show useful correlations to psychological and situative variables, but on the other hand we cannot find a trace of any influence; all deviations are in agreement with a pure chance behaviour (i.e. chance fluctuations).

NEW SYSTEMTHEORETICAL PARADIGMS FOR PARAPSYCHOLOGY

At the first glance the concept of influence seems to be so fundamental and so general that one can hardly imagine an alternative concept. System theory, however, could specify what is meant by the hitherto intuitive concept. From this point of view it is easier to find alternative concepts. During the last years system theory has investigated more and more complex, open and self organizing systems. It turned out that many new concepts were necessary to describe such systems, thus the boundaries of the familiar classical concept became obvious. If we assume that paranormal phenomena are a function of the human psyche as it is expressed in the standard definition, it is not farfetched to say that paranormal phenomena occur in the most complex system we know. Therefore we should have a look at the way normal science deals with such problems.

A central concept of modern system theory is the concept of autonomy. Nobody would doubt that the human psyche is at least to some extent autonomous. But even machines can be autonomous. Anybody who has ever worked at a computer terminal may have had this impression. But certainly it is necessary to define the term more precisely. F.J. Varela puts forward the following thesis (Varela, 1981): "Every autonomous system is organizationally closed." "An organizationally closed unity is defined as a composite unity by a network of interactions of components that (i) through their interactions recursively generate the network of interactions that produced them, and (ii) realize the network as a unity in the space in which the components exist by constituting and specifying the unity's boundaries as a cleavage from the background." After having discussed this definition in more detail, which we will not do here, however, he writes: "Autonomy means, literally, self-law. To see what this entails, it is easier to contrast it with its mirror image, allonomy, or external law. This is, of course, what we call control. These two images, autonomy and control, do a continous dance. One represents generation, internal regulation, assertion of one's own identity: definition from inside. The other represents consumption, input and output, assertion of the identity of other: definition from outside... The fundamental paradigm of our interactions with a control system is instructions, and the unsatisfactory results, errors. The fundamental paradigm of our interactions with an autonomous system is conversation, and its unsatisfactory results breaches of understanding."

These statements may intuitively sound familiar to those parapsychologists who know Stanford's conformance behaviour model or the recent developments of the observational theories. But at the first glance it is not clear how they can be applied to psychokinesis, for instance. As we have already said it is useful to assume that the human psyche (whatever this may be) could be regarded as an autonomous system in terms of the given definition. From the point of view of certain interpretations of quantum physics it is legitimate, to some extent, to consider quantum processes as autonomous too. But how could we bring both together? The simple assumption of Walker's model (Walker, 1975) that the non-local hidden variables which control the 'collapse of the wave vector' has been criticized by ourselves on several occasions (Lucadou and Kornwachs, 1977) and recently by P. Phillips (1984). Nevertheless we should not reject this idea totally; it may be considered as a first approximation. In a recent paper (Lucadou and Kornwachs, 1983) we tried to show how this model could be improved, and I cannot repeat the whole argumentation here. The essential point is that Walker identifies information rates processed in the brain with the information which is necessary to collapse the wave function during the measurement process. The problem is that he uses Shannon's classical concept of information. As we have stressed above, this implies the concept of influence. Such classical influence or information, howver, cannot be non-local as required in the hidden variable theory of quantum physics adopted by Walker. Such an assumption would contradict Lorentz's invariance of QM. Nevertheless the existence of non-local correlations in quantum physics remains and Walker is correct in assuming that they have some common features with psi phenomena. To solve the problem we have proposed a non-classical concept of information called 'pragmatic information'. Mathematically it is not a simple scalar function like Shannon's information, but an operator like other quantum physical observables, for instance like the projection operator which describes a measurement. The basic axioms of 'pragmatic information' are described in this paper. Here it is sufficient to say that they do not specify the amount of 'influence' which can be 'transmitted', but even in the case where this amount is zero, a non-local correlation could be obtained. The internal view of an autonomous system, however, as specified by Varela, fits exactly with the notion of correlation. Correlation is like conversation; we cannot find out the transmitter and the receiver, the part that influences and the part that will be influenced. Both are matched together; if we try to separate them the autonomy of the system will break down. Thus if we use the outside view and try to find out whether the subject has influenced the PK machine or vice versa, we would not get an answer. The only thing we can see is the correlation. This feature also fits with an uncertainty principle formulated for pragmatic information: We cannot describe a system simultaneously from inside and outside with arbitrary precision. In this context 'inside' means the structure of the system and 'outside' its behaviour (output).

What we have found here from a theoretical point of view is relevant in the practice of parapsychology. In his recommendations to experimenters in parapsychology Stanford (1974) points out that the subject should have the feeling that he or she "can handle the matter at its discretion or according to its own disposition or nature, and ... thus appropriately reduces one's level of concern and ceases to focus attention on the problem as such. Additionally, the cognition that agencies ... are already working and producing effects can serve to release PK effects in an individual." Now we understand why the usual instruction given to the subject "to influence the random generator" need not be optimal. So, why should we not say: "Please try to do PK!" But there is a further system theoretical reason: Autonomous systems are always self referential which is a feature of organizational closure. This can be expressed mathematically by the following equation:

$$J\,x(k) = j \cdot x(k) \qquad (3)$$

here $x(k)$ describes a state of the self referential system, and J an operator (e.g. matrix) which can be an observable of the system. The equation shows that a certain property of the system can be mapped within the same system. Very often it is argued that self reference would lead to a vicious circle, to a regress ad infinitum. But this is not true at all. Such equations are well known and fundamental in quantum physics. The Schrödinger equation has this form. The solution of such equations yields so-called eigenfunctions and eigenvalues. The special structure of this equation is responsible for all the well known 'paradoxes' of quantum physics. Now it turns out that the same holds for autonomous systems. (This also strongly supports our assumption that pragmatic information must be an operator and not a scalar function.)

Many popular science writers and unfortunately also many parapsychologists flatter themselves in the belief that quantum theory does contain an element of irrationalism. So they argue that quantum theory or the so-called 'new physics' shows that nature has a non-physical, non-causal background. But this conclusion is neither new nor correct. From the beginning of the development of quantum theory such ideas caused vigorous debates, as for instance F. Selleri points out in a historical overview (Selleri, 1983). But the concepts of system theory were not yet available, thus it is understandable that for instance the concept of causality and determinism were mixed up. We have discussed this distinction elsewhere. C.G. Jung as a genial precursor of the model we are discussing here, still called the new principle non-causal (Jung and Pauli, 1952), but his concept of synchronicity is very similar to our concept of non-local correlations. The original idea of Jung was to find a quasi mathematical dsecription of those processes which are attributed to what we call psyche. Of course at that time the formalism of system theory was not yet available but there is for instance an interesting relationship between his synchronicity theory and the theory of numbers (see Franz, 1970).

Furthermore, the so-called 'quantum paradoxes' do not indicate that there is any element of irrationalism within quantum theory. Quantum theory is logically and mathematically consistent and one of the best proven theories we ever had in science. The paradoxes are caused by our inability to 'think' in non-classical concepts. All the different 'interpretations of quantum theory' are merely the attempts to explain something progressive by something reactionary as C.F. v. Weizsäcker has formulated it (Weizsäcker, 1976). It seems highly probable that our inability to deal with the paranormal has similar grounds.

If we are encouraged enough to jump over this gap of classical 'thinking' (and today many scientists are prone to do this, see Jahn, 1984) we could define psi in the following way: Psi effects are (non-local) correlations within psychophysical systems which emerge from the autonomy (and/or autology) complexity and openness of those systems. Such 'non-classical' systems produce their own 'meaningful content' or 'eigenbehaviour' out of 'noise' which is only globally in agreement with classical thermodynamics (e.g. the second law). The subject and under certain conditions, even the observer, must be considered as a part of the (physical and psychological) system which is defined operationally by a given experimental design.

Until now we have not discussed the role of thermodynamics. Every macroscopic deviation from the equilibrium seems to violate the second law of thermodynamics. But this holds only for closed systems. Every living cell, however, is so far from the thermodynamic equilibrium that the probability for its existence would be near to zero. If life were only the result of a closed system it could not exist. But natural systems are not closed. (This is wrong in those calculations which start with the assumption that living cells are the product of pure chance, see Eigen, 1983). Thus I doubt whether the calculations of thermal noise model of R.D. Mattuck produces accurate estimates for a possible PK effect (Mattuck, 1979). However, this approach might be the first step in the right direction. The essential problem will be to find the correct variable which describes the openess of the system in question. Similar to Walker's model Mattuck starts with the assumption of a flow of information, which, however, from our point of view is not useful. One should not confuse thermodynamical openess with organizational closure. The former is a necessary condition of the latter. The view from 'outside', however, is important too: the overall behaviour of open systems is limited by the global flow of energy which dissipates in these systems. This yields the global limitations of all autonomous effects. Experience in parapsychology tells us that the effects cannot be too great (we have called this elsewhere "Timm's rule"). Another hint from this assumption is an effect which was already known to C.G. Jung: if we accumulate too many statistical experiments, every paranormal effect will be 'washed out'. In the paper mentioned above we have pointed out that this effect can also be explained by the properties of pragmatic information. Thus we recognize that Rhine's assumption that psi can be accumulated by statistical means must be limited.

Finally we have to clarify what is meant with the term eigenbehaviour. Above we noticed that self referential systems lead to a specific mathematical structure which is similar to the mathematical structure of quantum theory. Equation (3) produces eigenfunctions and eigenvalues. In relation to macroscopic systems these eigenfunctions are called 'eigenbehaviour'. The corresponding eigenvalues can serve as ordering parameters. The prefix 'eigen' is a German word and means 'self'. Thus eigenbehaviour is exactly that feature which is required in the definition of organizational closure. It will not be produced from outside the system but on the contrary it will be stabilized by the system itself against external influences. For living systems Varela et. al. called this feature 'autopoiesis'. Another important property of these eigenfunctions is that they can be discrete, they exhibit a quantum character. The perception of a 'gestalt' is one of the best known examples in psychology. It is a self stabilizing function and its quantum character is known as 'gestalt switch'. Furtheron 'perception of gestalt' describes a process where 'meaning' will be created from the context, but as we know, this meaning need not be unique. This is an important difference from L. Gatlin's model (Gatlin, 1977) of 'meaningful information creation' which starts from the ontological assumption of a totally deterministic world. In this model a real 'gestalt switch' could not occur without a given external influence.

On the other hand we recognize that from this point of view it is no longer possible to abstract any psi effect from its (meaningful) content. Moreover the 'meaning' can be regarded as that 'carrier' of the effect and exhibits its holistic caracter. The meaning of a special experimental setting, however, cannot be measured in units of Shannon-type information or in units of an influence but in units of pragmatic information.

Indeed we do not have the impression that psi is a self stabilizing function, but this is not in contradiction to our proposed definition. The existence of a stable eigenfunction does not imply that it must be observable directly, observability of the system need not be given at all, or, which is more likely in parapsychology, the observation process could disturb the system in such a way that the eigenstate in question will be destroyed. Such behaviour can be expressed in uncertainty relations which we have described in the other paper. Obviously such uncertainty relations confine any effect, but the degree of confinement can depend from the methods of measurement. In a paper on theoretical problems of RSPK we have stressed that, for instance, it may be advantageous to observe those phenomena deliberately with a reduced resolution (Lucadou, 1982).

Finally let me make some short remarks on the problem of learning again. Varela et. al. started their consideration from the point of view of biology, but there is another approach with very similar results which has its origin in a most abstract discipline, namely mathematical logic. In an early paper we have stressed how some results of this field could have an important impact on parapsychology. In a recent paper L. Löfgren applies these logical tools to the problem of learning (Löfgren, 1982). He writes: "Learning is one of those autological concepts, i.e., concepts that can be applied to themselves. We may for example learn to learn - which is fortunate, for if we could not, how could we then go about learning at all? Again, we may learn about learning - or else our learning ability would be obviously incomplete!" His use of the term 'autology' is very similar to the term 'autonomy' but it pronounces the logical point of view. Lofgren writes: "With autology we refer to the metalogics of self applicability, with foundational results in the work of Gödel, Tarski, Church, Kleene, Scott, and others. ... A characteristic result of autology is the so-called linguistic complementarity: descriptions and interpretations of a language are complementary. That is, as long as we stay within a language L, we cannot completely describe L only in terms of its sentences - both sentences and interpretations are required for a full account of L." (Again this is the same feature which Varela called definition from inside and from outside, or what we would call the complementarity between structure and behaviour (or function).) As a next step Lofgren proves that describable yet non-learnable effects may exist, supposing that learning is defined in terms of inductive inference which means 'rules which allow prediction'. In the section where we discussed the classical notion of influence we also noticed that control is an essential feature of learning. Here we learn that inductive inference is the logical counterpart of control. Now, what does this suggest for parapsychology? Obviously, the autological nature of learning goes beyond effectiveness or inductive inference. Thus we must be aware that even if we could not learn to produce psi effects in terms of control, there might be other unknown possibilities. The term 'experience' for instance might be better describe what is going on during a psi event. Thus it might not be accurate to speak of 'psi abilities', but instead of a spontaneous disposition or as Stanford does: 'conformance behaviour'.

At the moment, however, from all we have discussed here it is rather unlikely that we could solve the problem of psi by the use of plausible but far too simple models underlying the standard definition of psi. The new definition proposed above may still sound very unfamiliar and vague. But the standard definition is not less vague; it really does not say anything at all. The new concepts we use are not developed in parapsychology but in normal sciences where they have shown such a tremendous creative power that some scientists speak of a scientific revolution (Prigogine, 1979). There is no reason to believe that the disturbing anomalies which we call psi indicate the end of the scientific method.

ABSTRACT

The usual definitions of psi i.e. ESP and PK are considered more closely. The problems arising from these negative definitions are called the exclusion problem, the problem of preliminarity, the problem of the boundary conditions, the problem of the phenomenological consistency, and finally the problem of coherence of the describing language. These problems are already known in principle.

Many parapsychologists, being aware of the problems mentioned above, maintain to have an alternative, positive, yet intuitive concept of psi in terms of an undiscovered force or energy or information which emerges from psychophysical systems. It is argued, however, these these concepts contain the same implicit models as the usual negative definitions. Even the invention of the so-called operational definitions which only describe the experimental setting cannot avoid such implicit models. They enter for instance into the instruction given to the subject. The Rhinean paradigm is a good target to elucidate these hidden assumptions: (a) Psi is considered as an influence, which means an energy or an information transfer of an 'entity' (going out); (b) Psi can be accumulated by statistics or repeated observation; (c) Psi is considered to be ergodic which means that the result of an experiment is independent of the location and date (time) if all other conditions remain constant; and (d) A psi effect can be separated in principle from its content or meaning. Even though these inplicit models seem to be obvious or plausible by their simplicity and their general validity they might be wrong. Nevertheless such incorrect fundamental assumptions might prevent us from discovering the real structure of psi and thus force us to be confused by its elusive and chaotic character which is apparently in contrast with such simple models.

The concept of influence implies several other concepts such as information, control, observability, reachability, stability, and learning. These concepts are discussed in more details. It is shown how general systems theory is able to specify these terms and how empirical parapsychological findings in general contradict these specifications. In addition to these more general considerations, a special experiment is discussed which was performed at Freiburg University with the partial aim to test some of these implicit models discussed above. The experiment can be regarded as a successful PK experiment under the usual parapsychological paradigm, nevertheless it clearly indicates that some of the basic assumptions mentioned above must be wrong. It especially indicates that there is no 'influence' on the radioactive source of a Schmidt-type PK machine which means that there is no change in the decay-rate of the source. Further, it seems to be highly probable that there is no 'influence' on any other part of the machine in the sense that an 'incoming' information or an 'influx' has produced the statistical deviations. It seems that these deviations are in agreement with the stochastical overall behaviour of the random generator. Nevertheless, significant correlations to psychological or situational variables could be found. Thus we are forced to introduce new fundamental paradigms to describe the empirical findings of parapsychology more adequately.

The concepts which are associated with complex, open, and self organizing systems such as autonomy, organizational closure, and correlation are discussed. From this point of view the basic assumptions of the observational theories are criticized and improvements are proposed. The concept of pragmatic information as a non-classical alternative to Shannon's concept of information is shown to represent typical features of autonomous systems. The relation to other models such as the conformance behaviour model or synchronicity model are mentioned. The new system theoretical paradigms imply features of wholeness, non-locality, and uncertainty. These considerations lead to the following suggestions towards the definition of psi:

'Psi effects are (non-local) correlations within psychophysical systems which emerge from the autonomy (and/or autology), complexity, and openness of those systems'

Such 'non-classical' systems produce their own 'meaningful' content or 'eigenbehaviour' out of 'noise' which is only globally in agreement with classical thermodynamics (e.g. the second law). The subject and the observer must both be considered as a part of the system which does not mean, however, that boundaries of the system cannot be given.

Such correlations or 'psi effects' can only be validated post hoc within classical models which means that predictions cannot be given on the basis of a simple scalar function. This implies strong limiting principles for any psi effect which is described elsewhere. One should add, however, that this model is quite different from L. Gatlin's 'meaningful information creation' model which starts with the ontological assumption of a totally deterministic world. Further, it is discussed whether we could still consider psi to be a learnable ability . System theory has shown that describable yet non-learnable events can occur. The given definition is now a positive one as is usual in natural sciences which can be applied in experimental work.

REFERENCES

Bauer, E. and Lucadou, W.v. 'Methoden und Ergebnisse der Psychokinese-Forschung', in Condrau, C. (ed): Transzendenz, Imagination und Kreativität. Die Psychologie des 20. Jahrhunderts. Band XV, Kindler, Zürich 1979, 494-512.

Eigen, M. 'Entstehung des Lebens', Natur, 1983, 3, 68-77.

Franz, M.L. 'Zahl und Zeit', Stuttgart, Klett, 1970.

Gatlin, L. 'Meaningful information creation: An alternative interpretation of psi phenomenon', J. A.S.P.R., 1977, 71, 1-18.

Hövelmann, G.H. 'Zum Problem der Wiederholbarkeit parapsychologischer Experimente', Z. f. Parapsych. u. Grenzgeb. Psychol., 1983, 25, 29-54.

Jahn, R.G. 'The persistent paradox of psychic phenomena: An engineering perspective', Proc. of the IEEE, 70, 2, 136-170.

Jahn, R.G. and Dunne, B.J. 'On the quantum mechanics of consciousness, with application to anamolous phenomena', Princeton, N.J., Princeton University, 1984.

Jung, C.G. and Pauli, W. 'Naturerklärung und Psyche', Zurich, Rascher, 1952.

Kliemann, W. and Müller, N. 'Logik und Mathematik für Sozialwissenschaftler 2', Munchen, UTB W. Fink, 1976.

Löfgren, L. 'Excerpts from the autology of learning', in R.Trappel (ed): Cybernetics and systems research, Amsterdam/New York/Oxford, North Holland, 1982.

Lucadou, W.v. and Kornwachs, K. 'Can quantum theory explain paranormal phenomena?', in E. Morris (ed): Research in parapsychology, 1976, Metuchen, Scarecrow Press, 1977, 187-191.

Lucadou, W.v. 'A new computer-controlled device for testing different PK-hypotheses', in W.G. Roll et. al. (eds): Research in parapsychology 1982, Metuchen/London, Scarecrow Press, 1983, 165-168.

Lucadou, W.v. 'Der flüchtige Spuk', Z. Parapsychol. u. Grenzgeb. Psychol., 1982, 24,

Lucadou, W.v. and Kornwachs, K. 'Beyond reductionism - reconsidered', E.J.P., 1983, 4, 357-376.

Lucadou, W.v. and Kornwachs, K. 'On the limitations of psi', Psychoenergetics, 1983, 5, 53-72.

Mabbett, I.W. 'Defining the paranormal', J.o.P., 1982, 46, 4, 337-354.

Mattuck, R.D. and Walker, E.H. 'The action of consciousness on matter: A quantum mechanical theory of psychokinesis', in A. Puharich (ed): The Iceland papers, Amherst, Essentia Research Associates, 1979.

Phillips, P.R. 'Measurement in quantum mechanics', J. S.P.R., 1984, 52, 797, 297-306.

Prigogine, I. 'Vom Sein zum Werden', Munchen, Piper, 1979.

Schmidt, H. 'Instrumentation in the parapsychology laboratory', in J. Beloff (ed): New directions in parapsychology, London, Elek, 1974, 13-37.

Selleri, F. 'Die Debatte um die Quantentheorie', Braunschweig/Wiesbaden, Vieweg, 1983.

Stanford, R.G. 'An experimentally testable model for spontaneous psi events', J. A.S.P.R., 1974, 68, 34-57 and 321-356.

Stanford, R.G. 'An experimentally testable model for spontaneous psi events', J. A.S.P.R., 1974, 68, 34-57 and 321-356.

Varela, F.J. 'Autonomy and autopoiesis', in G. Roth and H. Schwengler (eds): Self-organizing systems, Frankfurt/New York, Campus, 1981, 14-23.

Walker, E.H. 'Foundation of paraphysical and parapsychological phenomena', in L. Oteri (ed): Quantum physics and parapsychology, New York, Parapsychology Foundation, 1975, 1-53.

Weizsäcker, C.F. 'Personal communication, 1976.

Wolman, B. B. (ed) 'Handbook of parapsychology, New York, Van Nostrand Reinhold, 1977.

Walter von Lucadou,
Hildastrasse 64,
D-7800 Freiburg,
W-Germany

EUR.J.PARAPSYCHOL., V.5, 1984

ARE PSI EXPERIMENTS REPEATABLE?
A CONCEPTUAL FRAMEWORK FOR THE DISCUSSION OF REPEATABILITY

Gerd H. Hövelmann

INTRODUCTION: A THREEFOLD CONFUSION

At least since the early days of regular experimental laboratory research in parapsychology, researchers engaged in this field and their critics alike have commonly considered the question of repeatability of parapsychological experiments one of the most momentous problems with regard to the future of this field of inquiry. However, if one attempts to survey these discussions about the repeatability problem in parapsychology during the last half of a century (note 1), as I have done in some detail elsewhere (Hövelmann, 1983b), it soon becomes quite obvious that the respective positions on this matter taken among parapsychologists (as well as those taken among critics) are much less uniform than one might have expected. The entire range of possible views of this problem and its importance to parapsychology, which may come to mind, is, in fact, to be found in the parapsychological literature. Thus, at the one extreme, Tart, Puthoff, and Targ (1979), the joint editors of 'Mind at Large', in the introduction of their book assert that experimentally produced psi phenomena are "robust and hidden" (ibid., p.xiv), and they claim that "we have finally come close to having repeatable psi experiments" (ibid., p.xv). In his review of this book, Stokes (1980, p.445) has justly criticized these exaggerated assessments as "propagandistic overselling of parapsychological research", and he has complained that the authors seem to be looking at parapsychological research "through rose-colored glasses" (ibid.). At the other extreme on the scale of assessments of repeatability of psi experiments, we find Blackmore's (1983) recent conclusion that parapsychology will only survive as a scientific endeavour, if it accepts the 'nonrepeatability of psi' and realizes that it needs a new hypothesis. Therefore, she recommends that we start all over again with the same phenomena, but a different hypothesis. A great number of various intermediate positions are to be found in the parapsychological literature (note 2).

A second fact rendering a solution to the repeatability problem all the more difficult is the vague and diffuse way that the problem is usually formulated; such terms as 'repeatability', 'replicability', and 'reproduceability' are frequently used interchangeably and in non-uniform and highly ambiguous ways. The fact that some authors use the selfsame terms to talk about the repeatability of experiments which are used by others for talking about replicability of effects, further adds to the confusion.

As I observed in my review (Hövelmann, 1983a, p.74) of the Parapsychology Foundation's proceedings of a conference on 'Parapsychology and the Experimental Method' (Shapin and Coly, Eds., 1982), a third and maybe the most important obstacle in the way of an adequate treatment of and a final solution to the repeatability problem appears to be a lack of proper understanding among parapsychologists as well as among their critics of what an experiment is. Researchers in parapsychology (as well as in many other empirical disciplines) too readily content themselves with a rather confused comprehension of the nature and purposes of experimentation. Naturalistic misconceptions of the purposes of experiments are the order of the day. Obviously, such shortcomings cause little or no problems to other empirical sciences because at least some of these sciences dispose of sufficiently repeatable experiments. However, the situation is quite different in parapsychology where, for the very reason of lacking repeatability, we are badly in need of an appropriate understanding of 'experiment'. Unfortunately the question as to what makes an experiment an experiment was not adequately addressed at the aforementioned conference on 'Parapsychology and the Experimental Method'. The same seems to be true, to judge from the abstracts of the presented papers (Beloff et al., 1983), of the 32nd International Conference of the Parapsychology Foundation on 'The Repeatability Problem in Parapsychology', which was held in San Antonio in October of 1983.

Therefore, it may be useful here to examine this question in some detail and to try to provide a conceptual framework for future discussions about the repeatability problem. I am confident that this can be helpful to avoid the threefold confusion which I have briefly sketched above. I will proceed in the following steps: first, I will introduce a systematically justifiable (note 3) conception of 'experiment' and of 'repeatability of experiments', which makes use of a constructive theory of action and which can be shown to be appropriate for the so-called natural sciences (at least for physics and chemistry). Then, along with several terminological clarifications and suggestions, I will explicitly specify the conditions which are suitable to guarantee repeatability of experiments in these disciplines. The application of this understanding of 'experiment' and 'repeatability of experiments' to any of the other scientific branches is subject to certain more or less decisive methodological (and pragmatic) restrictions, and I will try to point out some of these restrictions which apply in the case of parapsychology.

THE OPERATIONAL BASIS OF EXPERIMENTS: ACTIVE INTERVENTION

Let us start with the following questions: What is an experiment? And why is it possible to gain empirical knowledge by way of experimentation? These seem to be rather simple questions at first sight, but on closer inspection it becomes obvious that they require not-so-simple answers. If we ask a natural scientist, as for instance a physicist who has not yet been rattled by philosophy of science, to answer these questions, we may expect one or both of the following naturalistic justifications of experiments in reply, which largely correspond with current textbook opinion:

a) Experiments are performed to gain empirical knowledge about events which seem to be contingent. That is, experiments are means to make certain events appear regularly, which previously and at irregular intervals have been observed in nature ('In order to study the law of falling bodies, one needs not wait for windfall'). In current textbooks of physics, then, formulations like the following are frequently to be found: we are told, for instance, that experiments provide opportunities to 'put questions to nature', to 'find out what it is that holds the world together' (note 4), to 'discern reality', or (as some natural scientists preferably express themselves) to 'eaves-drop nature's secrets' (cf. Höfling, 1971, pp.18, 233, 331, 332, passim). For detailed criticisms of these ways of speaking about experiments, see Janich (1977) and Hövelmann (in press - d).

(b) Experiments are performed to gain empirical knowledge about nature under 'pure' conditions. That is to say, that in experiments certain idealizations are unavoidable since, according to our fictive interlocuter, natural processes are too 'intricate' (Höfling, 1971, p.17), too 'colorful' (ibid., p.330), and too 'multiform' (ibid., p.329) for the 'human intellect' to grasp their 'real' substantiality (ibid.). Therefore, to give two examples, we are told that, in chemical experiments, substances of such a purity are used as they are not to be found anywhere in nature or that physicists are studying motions on an inclined plane in vacua, at first, in order to be able to disregard possible influences by friction, size, and inert mass of the body to be investigated, or other parameters. By means of such idealizations, Höfling (1971, p.329) asserts, the "human intellect is enabled to follow natural processes" (note 5).

As a consequence of inadmissably curtailed conceptions of science, as they are held, for instance, by Logical Positivists of the Vienna Circle (cf. Hempel, 1966) and Critical Rationalists of the Popperian school of philosophy of science (cf. Popper, 1935/1959, 1979; Albert, 1980), a fundamental distinction is completely ignored in these answers. This neglected aspect is the fact that only the actions of the human being, of the scientist, of the 'homo faber' render science and scientific experience possible and form the pragmatic as well as the justificatory basis (cf. Hövelmann, in press - c) for scientific knowledge. Human action plays a very decisive role in the production of scientific knowledge; that is, it constitutes scientific knowledge, first of all. Human actions, which are the basic preconditions or - to use Kant's famous phrase - "the conditions of the possibility" of gaining scientific experience and scientific knowledge, are completely lost sight of in the aforementioned answers from our fictive, but representative interlocuter. In these typical answers, it is pretended that scientists only become active in order to make those events and processes easier to observe which, irrespective of and prior to any experimental action, are believed to be happening in nature anyway. But how, one wonders, should it actually be possible to substantiate the claim that the same events and processes, which we study in the laboratory, are also and independently happening in (undisturbed) nature, when this claim itself is obviously based on experiments which use technical tools, prototypes of which are not at all to be found in nature? (cf. Tetens, 1982).

Results of experimental actions, moreover, are only understandable and interpretable to the purposes of the respective actions. Actions are always means to certain ends. If someone does not know anything about the ends, which he is trying to attain by means of certain actions, he is unable to say anything about the results (failures or successes) of these actions, afterwards. Analyzable results can only be obtained by way of active interventions in given situations with the intent to bring about another situation. It can only be judged against such a different situation whether the performed active interventions really were appropriate means for the attainment of the respective ends. These ends, moreover, have to be explicitly specified and formulated in advance. The construction, building, and usage of experimental tools which, likewise, are understandable only by a consideration of the intentions and purposes the experimenter associates with them, are also to be regarded as active interventions in given situations. Thus, knowledge about successes and failures of actions can only be reached by judging the results of these actions against their purposes, which, to repeat this, have to be explicitly formulated in advance. It follows that scientists are only able to learn something new by way of experimentation (that is, an experiment is only informative), if a particular result of the experimental actions has been predicted. This 'instrumentalist' knowledge, which is reached by way of an active intervention into a given situation, is in marked contrast to 'contemplative' knowledge (Janich, 1981, p.424), which can be reached by merely passively observing a situation (note 6).

The everyday, pre-theoretical conception of 'experience' already implies that, as an acting person, one can succeed or fail in one's respective purposes. Thus, experience can only be gained by reaching or missing one's purposes. The difference between everyday and scientific experience is merely due to more rigorous standards required for the latter.

On the basis of the fundamental human faculties (Dingler, 1955, p.16; Lorenz and Mittelstrass, 1967) to speak, to act, and to speak about actions, it is possible to draw a distinction between 'natural' and 'artificial': those things and events, which are to be found in nature (in the broadest sense) and - even more important - which are not (or not yet) manipulated by man, are to be called 'natural'; those things and events, which - directly or indirectly - result from man's actions, on the other hand, are to be called 'artificial'. Tools and events, which man has set into operation (and, of course, this includes experiments!), are examples of artificial things and events. In the so-called natural sciences, almost every investigation makes use of experiments and of (technical) tools (of artifacts, that is). Therefore, to follow Janich's (1980, p.87) suggestion, one should perhaps better speak of 'Experimentalwissenschaften' ('experimental sciences') instead of 'natural sciences', because these sciences are exclusively dealing with "artificial properties of tools and events, which are produced by man, and also with regularities which are artificially brought about and maintained" (ibid.; my translation).

I consider Janich's term 'experimental sciences' ('Experimentalwissenschaften') - Tetens and Janich also speak of 'technical sciences' ('Technikwissenschaften') - quite useful, and I will adopt it for the present paper. However, I will use this term in a way that slightly differs from Janich's usage. While Janich uses the term 'experimental science' as a substitute for 'natural science', it will be used here to denote a certain type of scientific activity that makes use of experiments, the objects of which do not act themselves. In contrast to 'experimental science' or 'experimental activity', a different type of scientific activity that also makes use of experiments, but the objects of which are (autonomous) human beings, who act and speak (or who are at least able to act and endowed to speak), will be called 'experimenting science' or 'experimenting activity' henceforth. Consequently, I will also distinguish between experimental and experimenting purposes, actions, effects, etc.. Obviously, this terminological distinction between certain types of scientific activities neither is a strict distinction between scientific branches, nor is it an equivalent to the distinction between the so-called 'natural sciences' and the 'cultural', 'behavioral', or 'social sciences': while in the classical 'natural sciences' like physics and chemistry, all activities which make use of experiments are 'experimental' activities (for convenience, physics and chemistry may thus be called 'experimental sciences'), this is not true for another of the natural sciences: biology. Activities in biology can be either 'experimental' or 'experimenting', depending on the type of experiment that is performed in a given case. The same applies to experiments in medicine and pharmacy. While experiments in psychology (note 7) and parapsychology (almost) exclusively require 'experimenting' activities, this is not true for a number of experiments in physiology. Obviously, a great number of scientific activities in the behavioral sciences (that is, those activities which do not involve experiments of any kind) are neither 'experimental' nor 'experimenting' activities. As will become evident as this paper proceeds, the distinction between experimental and experimenting activities and sciences cuts across the different scientific disciplines. That is, it is not so much a distinction between those scientific branches where strict repeatability is a possibility and those where it is not; rather, it is a distinction between those scientific activities which lead to strict repeatability and those which do not.

In experimental sciences, then, in which experimental actions lead to strictly repeatable experiments, the realms of 'the natural' is definitely forsaken, because, as was pointed out above, a repeatable experiment also is an artificially produced event. The criterion of repeatability that applies here is the possibility to artificially and systematically cause events (i.e., experimental effects), which would not regularly turn up in this way without man's active intervention. That is to say, that, in principle, these events are always at the experimenter's disposal. Thus, replicable quantitative experimental findings are not natural events enticed into the laboratory, but rather artificial products of experimental actions.

REPEATABILITY OF EXPERIMENTS IN THE EXPERIMENTAL SCIENCES

What do we mean when we claim that experiments in the experimental sciences are repeatable? What, precisely does it mean when we claim that strict repeatability of experiments is an essential characteristic, say, of physics?

Since in the case of a negative outcome (note 8) of an experiment, which was intended to replicate findings of an earlier experiment, one always must decide whether this experiment has, in fact, been a repetition of the earlier experiment, the definition of the essentials of repeatable experiments must not be based on the outcomes of the experiments. The outcome of an attempted replication cannot be a sufficient criterion of the repeatability of the experiment in question (note 9).

In experiments, not natural events (in the sense specified above), but only human actions as well as processes, states, and situations, which are directly brought about by human actions, can be repeated and scrutinized. From this it follows that the conditions under which certain outcomes come about (rather than these outcomes themselves) must be repeatable. In addition, the tools which are used in the experimental sciences to produce, demonstrate, and record these outcomes must be reproducible. To illustrate what I have said so far, I would like to introduce the following scheme of the standard form of experiments in the experimental sciences:

$$S_0 \quad a \quad S_1 \quad C \quad S_2 ,$$

where 'S_0', 'S_1', and 'S_2' stand for three different situations (0, 1, and 2), 'a' stands for one or more actions of an experimenter, and 'C' stands for a course, which does not involve any further action of an experimenter. This scheme is to be read: By means of an action or a sequence of actions (a) an experimenter transforms a given situation S_0 into the starting situation S_1 from which, by way of a course C, which does not involve any further active intervention on the part of the experimenter, the final situation S_2 results. To be able to realize the transformation of S_0 into S_1, the experimenter must follow the instruction $!a_{1,\ldots,n}$ (read: perform the actions $a_{1,\ldots,n}$!). If the experimenter conscientiously observes the instruction $!a_{1,\ldots,n}$, the actions $a_{1,\ldots,n}$ can guarantee that the purposeful and carefully directed transformation of S_0 into S_1 succeeds (note 10). For the formulation of these instructions, explicitly standardized and operationally well-defined terminological means are required (Hövelmann, 1983c, p.131; in press - a; in press - b).

The argumentation so far enables us to answer one of our main questions about the nature, the essentials, and the purposes of experiments: an experiment is the purposeful, systematic, and carefully directed application of a disposable 'know-how' (that is, knowledge about the feasibility and the immediate effects of our actions) with the intent to produce an otherwise non-disposable 'know-that' (that is, knowledge about courses), which comes about as a consequence of our active modifications of given situations and of our active preparation of experimental conditions in the situation S_1. Such a 'know-that' can only be reached, however, if we have formulated the purpose of our experimental investigation (or, synonymously, a hypothesis) in advance, because otherwise we will not be able to distinguish failures from successes (certain effects from a malfunctioning of our tools, that is). The experimenter acts to make certain arrangements (construction of measuring tools, experimental set-up, etc.) by means of which a given situation S_0 is transformed into the situation S_1, which represents the starting situation for a course. 'Repeatability of an experiment' in an experimental science, thus, always refers to the strict repeatability (i.e., the reconstructability) of the situation S_1 and the strict repeatability of the actions $a_{1,\ldots,n}$, which have led to S_1.

The equality of respective starting situations S_1 in subsequent experiments, which attempt to replicate findings of an earlier experiment, with the starting situation in that original experiment again depends on (the availability of) adequate terminological means. Completely equal starting situations in subsequent experiments can only be reached, if the starting situation in the original experiment as well as the experimental actions which produced this situation are explicitly and exhaustively described so that, analogous to the observance of a recipe (note 11), S_1 can always be exactly and reliably reconstructed. If there are doubts as to whether the experiment in question, in fact, is a repetition of the earlier experiment, it must be decided whether the actions $a_{1,\ldots,n}$, which have led to the situation S_1 , have been performed in strict adherence to the 'recipe' (to the instruction $!a_{1,\ldots,n}$, that is). It follows from all this, that statements about repeatability always refer to technical-experimental actions of experimenters, and never to 'lawfulness of nature'. It makes no sense to claim that nature is 'lawful', because everything we know about nature in the sense of empirical physics is knowledge about the experimental production of technical effects. In this sense, an experimental set-up in physics can be described as a machine for the production of such effects. What we are used to calling 'natural law' always refers to the pre-planned, technically enforced functioning of our experimental apparatus (that is why I prefer to speak of 'experimental laws' instead of 'natural laws'). There is only one point where nature (in the strict sense specified above) enters into a physical experiment: in the last resort, we are dependent on natural materials for the production of our experimental apparatus (note 12). And, inevitably, natural materials are more or less resistant to our manipulative, technical attempts to work them. This non-eliminable natural influence is reflected in the constant factors in our mathematical descriptions of tool functioning. These constants keep reminding us that in the pursuance of technical-experimental purposes we remain dependent on the properties of natural materials (note 13).

In the introduction, I mentioned that the terms 'repeatability', replicability', and 'reproduceability' are frequently used in non-uniform and ambiguous ways. Now, the reader will have realized that I have used all of these terms in the foregoing pages. After having explained what I take to be the standard form of experiments in the experimental sciences, I will now make terminological suggestions for the use of these terms which reflect their use in the present paper.

I suggest that the term 'repeatability' and its derivatives (to repeat, repetition, etc.) be reserved for talking about exactly defined experimental actions and about starting situations (S_1) for courses (C) which are brought about by these actions. Thus, it makes sense to speak of repeatable actions, of repeatable experimental conditions, and, generally, of repeatable experiments. Experimental effects or findings, however, may not be said to be repeatable. (Repeatable experimental conditions may also be called 'reconstructible' conditions).

The term 'replicability' and its derivatives (to replicate, replication, etc.) should be used to denote experimental effects (S_2) or the results of experiments. On the condition that experiment B is a repetition of experiment A (in the sense specified in the preceding paragraph), it makes sense to say that experiment B has replicated (or failed to replicate) the findings of experiment A. If, on the other hand, experiment B fails to repeat experiment A, we are unable to say anything about the replicability of the effect which was found in experiment A (even if the effect of B appears to be 'similar' to or of the same kind as that of A); in this case a different type of experiment has been performed.

The term 'reproduceability' and its derivatives (to reproduce, reproduction, etc.), finally, should be reserved for talking about the use (in the experimental set-up of experiment B) of exact copies (Tetens, 1977) of measuring tools, instruments, and pieces of apparatus, which has been used in the experimental set-up of experiment A. That means that neither experiments nor experimental findings may be said to be reproducible.

To summarize the argumentation so far, the following are the essentials of experiments in the experimental sciences: (1) By means of a purposeful, systematic, and carefully directed application of a disposable 'know-how', a given situation (S_0) is transformed into a different situation (S_1), which serves as the starting situation of a course (C). (2) This course, which does not involve any further active intervention on the part of the experimenter, results in a final situation (S_2). (3) In this way, an otherwise non-disposable 'know-that' (i.e., knowledge about courses) is obtained. (4) This 'know-that' systematically depends on the foregoing artificial production of the situation (S_1) by means of 'know-how', which is formulated in an appropriate instruction ($!a_{1,\ldots,n}$). (5) This shows that the objects of the experimental sciences are artificially produced effects rather than 'nature'. (6) Therefore, 'repeatability' always refers to human actions and to the situations which are immediately produced by these actions. (7) In order to guarantee repeatability, experimenters must follow instructions $!a_{1,\ldots,n}$ for the production of adequate starting situations for courses. (8) The purpose of an experiment (i.e., the hypothesis to be tested) must be explicitly formulated in advance; that means that a particular outcome of the experiment in the final situation (S_2) must be predicted, because otherwise it would be impossible to tell a successful from a failing experiment, an experimental effect from the malfunctioning of the apparatus. (9) To be able to formulate the instructions $!a_{1,\ldots,n}$ as well as the hypothesis to be tested and the predicted results, the experimenter must dispose of a standardized, methodically constructed, and operationally well-defined terminology, which is free of circular definitions. (10) The 'repeatability' of experiments can unequivocally be distinguished from the 'replicability' of experimental effects and from the 'reproduceability' of the experimental apparatus. (11) The repeatability of the starting situation S_1 and of the experimental actions $a_{1,\ldots,n}$ is a necessary precondition of the replicability of effects in the final situation S_2. One should only talk about the replicability of effects on the condition that the repeatability of the starting situation and of the actions, which produced this situation, is already established. (12) Empirical scientific knowledge (not exclusively, but) decisively depends on the actions which experimenters perform; experiments are not a passive reception of 'whatever nature may decide to tell us', but an active production of 'know-that' that systematically depends on our respective 'know-how'.

METHODOLOGICAL RESTRICTIONS

As I have observed above and as I will again emphasize below, all experiments of whatever kind are artificially produced events. Nevertheless, as I indicated in the introduction, there are important methodological differences between experimental and experimenting sciences, because the conceptual framework I developed in the first part of this paper for the discussion of repeatability in the experimental sciences is subject to a number of methodological restrictions when it is applied to experimenting sciences like psychology and parapsychology. In the case of parapsychology, we can, I think, distinguish between relative methodological restrictions, on the one hand, which are rooted in insufficiencies of our current practice or theories and which, in principle, may be remedied sooner or later, and absolute methodological restrictions, on the other, which spring from differences between experimental and experimenting scientific activities and which, therefore, can not be avoided. I will start with a consideration of two relative methodological restrictions, which apply in the case of parapsychology, and deal with absolute methodical restrictions afterwards. I will close this discussion with a list of those parts of the conceptual framework which, I think, can be preserved in parapsychology.

Doubtlessly, one of the most important relative methodological restrictions of the applicability of the conception I have developed above, is parapsychology's lack of adequate terminological means. As I have tried to show, a standerdized, methodologically constructed, operationally well-defined, and coherent terminology is indispensable for the formulation of experimental purposes, for the prediction of a particular outcome of the respective experiment, and for the formulation of the instruction $!a_{1,...,n}$. The obscure, inconsistent, and ambiguous terminological means, which presently are at our disposal, cannot be used for a clear, unmistakable, and unambiguous formulation of instructions, experimental purposes, etc. This problem can, in principle, be solved by the construction of an adequate terminology (cf. Hövelmann 1983c; in press - a, esp. section III.2). Very recently, several other parapsychologists have expressed growing concern about the inappropriate terminological means of the field; Carlos S. Alvarado, Vernon M. Neppe (in press), John Palmer (1984), Michael A. Thalbourne (1984), and Nancy L. Zingrone are among them. This recent development reinforces my hope that we will eventually get rid of this unpleasant problem.

The second relative methodological restriction of the applicability of the understanding of 'experiment' and of 'repeatability of experiments' in the experimental sciences to experiments in the experimenting sciences like parapsychology, which I want to discuss here, concerns the courses (C), which are to be investigated in experiments. As I explained above, a course (C) must not involve any active intervention on the part of the experimenter. The same must necessarily be postulated for experiments in the experimenting sciences, because otherwise it would be impossible to distinguish experimenting effects from immediate effects of actions (which do not need immediate courses). However, there is an additional problem in parapsychology: if there really is something that can be described as a psi-based experimenter influence (for instance, cf. Broughton, Millar, Beloff, and Wilson, 1978; Broughton, 1979; Millar, 1979), then it must be clarified whether this influence is to be regarded as an action or as a mere behaviour of an experimenter (note 14). In the latter case, this influence would form part of a course (C). If a psi-based experimenter influence is to be regarded as an active intervention on the part of the experimenter, however, then it must be considered a violation of the postulate that courses be free of active interventions of an experimenter.

To turn now to the absolute methodological restrictions, the following departures from the standard form of experiments in the experimental sciences seem unavoidable if experiments in the experimenting sciences are considered:

As I have observed above, the distinctive criterion that separates experiments in the experimental sciences from those in the experimenting sciences is the involvement of living human beings as the objects of experimentation in the latter sciences. This entails important consequences in more than one respect: (1) Parts of the actions $a_{1,\ldots,n}$, which an experimenter in the experimental sciences must perform in order to transform a given situation (S_0) into the starting situation (S_1), consist of the adjustment or regulation of his set of physical apparatuses. In the experimenting sciences, on the other hand, the experimenter is dealing instead (or in addition) with living organisms. What has to be 'adjusted' or 'regulated' here is such a living human (or animal) organism. Evidently, these subjects in experiments of the experimenting sciences cannot be 'switched on' and 'off' in the same way as a physical testing apparatus can be prepared for experimentation (in the present paper, I will not deal with the ethics of experimentation with human and/or animal subjects). (2) An experimenter in the experimental sciences 'acts on' certain inanimate physical materials, while the experimenter in the experimenting sciences 'interacts with' (cf. Hövelmann, 1984) living organisms. While experimental tools only do what an experimenter forces them to do, living subjects need not remain passive, but can, in principle, do what they choose to do. (3) The psychological dispositions of experimenter and subject and various other situational and motivational variables can greatly influence (and even prevent) the successful production of a starting situation S_1. The psychological dispositions of the subject, moreover, may (and in many cases is likely or expected to) influence the final situation S_2. (4) In the experimenting sciences (just as in the experimental sciences), the course (C) must not involve an active intervention of the experimenter, but it may well involve (quite contrary to the situation in the experimental sciences) one or more actions of the subjects of the experiment (and, by definition, human actions can never be part of 'know-that'; cf. Hövelmann, 1984, pp. 145-152). (5) From (1) and (4), it follows that an additional set of instructions ($!A_{1,\ldots,n}$) is needed in most of the experiments in the experimenting sciences to tell the subjects how they are supposed to act in the experiment. (6) If we disregard animal experiments for a moment, we can say that the subjects in experiments of the experimenting sciences are living human beings, who are acting and speaking (or who are at least able to act and are endowed to speak) and who are, moreover, able to learn by means of action and speech. Now, the experimental instructions to the subjects ($!A_{1,\ldots,n}$) are necessarily linguistically composed. It is quite obvious, therefore, that the subject's comprehension of these experimental instructions (no matter whether written, tape-recorded, or other instructions are used) always depends on his or her individual life history and individual learning history. (7) Again, because of the human ability to learn, it is impossible to confront a human being with the same situation twice. (8) The success of the experimenter's 'know-how' partly depends on the actions and behaviour of the subject involved in situation S_1. (9) Even if experimenters dispose of adequate terminological means, they are unable to make any exact predictions about a particular outcome in the final situation S_2. All the experimenters can do is to predict that, with a certain degree of probability, the courses $C_{m_1},\ldots,C_{m_n}$, which start in situations $S_{1m_1},\ldots,S_{1mn}$, will finally result in situations $S_{2m_1},\ldots,S_{2mn}$. (10) While the experimenting actions $a_{1,\ldots,n}$ must also be repeatable in the experimenting sciences, this is not necessarily true for the starting situation S_1. It can not be guaranteed here that strict observance of the instructions $!a_{1,\ldots,n}$ determines exactly any single starting situation S_1. Thus the above experimental-scientific postulate must be somewhat 'liberalized' to allow for the production of a range ($m_{1,\ldots,n}$) of possible starting situations $S_{1m_1},\ldots,S_{1mn}$ by means of the actions $a_{1,\ldots,n}$ (under Point 9, we have already made use of this fact). These absolute methodological restrictions result in some important modifications of the standard form of experiments in the experimental sciences. A standard form of experiments in the experimenting sciences may then be schematized in the following way:

$$
\begin{array}{cccccc}
 & & p(S_{1m_1}) & C_{m_1} & p(S_{1m_1}) \\
 & & \vdots & \vdots & \vdots \\
S_0 & a & \vdots & \vdots & \vdots \\
 & & \vdots & \vdots & \vdots \\
 & & p(S_{1m_n}) & C_{m_n} & p(S_{1m_n})
\end{array}
$$

From this, it follows that the (still frequently advanced) postulate that experiments in the experimenting sciences (and in parapsychology in particular) should be strictly repeatable, cannot be substantiated in a methodologically justifiable way.

Finally, I will briefly list those aspects of experiments in the experimental sciences, which can be preserved in the experimenting science (in order to avoid possible confusion, I may remind readers once more that, according to my terminological distinctions, the terms 'experimental' and 'experimenting' sciences are not equivalent to the terms 'natural' and 'social' sciences; they rather denote different types of scientific activity):
(1) Results of experiments decisively depend on the actions the experimenters perform. (2) Experiments are dealing with artificially produced situations and effects, but not with 'nature'. (3) A standardized, methodologically constructed terminology is obligatory. (4) The experimenter has to adhere to well-defined instructions for the production of (a range ($m_{1,\ldots,n}$) of) starting situations. (5) 'Repeatability' always refers to human actions and to the situations which are immediately produced by these actions. For experiments in the experimenting sciences, this means that the starting situations which has actually been produced by these actions $a_{1,\ldots,n}$, must repeatably be one of a range ($m_{1,\ldots,n}$) of several possible starting situations. (6) A range ($m_{1,\ldots,n}$) of possible final situations must always be explicitly formulated and predicted in advance. (7) 'Repeatability' of experiments can unequivocally be distinguished from the 'replicability' of experimenting effects and from the 'reproduceability' of the experimenting apparatus. Thus, the aforementioned definitions of 'repeatability', 'replicability', and 'reproduceability' can be retained.

To conclude, the present study probably has not moved us a single step further towards a solution to the repeatibility problem. Such a solution was beyond the purpose of this paper, after all. However, my conceptual and methodological considerations and differentiations have provided a framework which, I hope, will contribute to easier, more fruitful, better organized, and less confused future discussions about the perennial problem of repeatability of parapsychological experiments (and of those of other experimenting sciences). In addition, they provide systematically justifiable arguments against the demand, still advanced by a few critics, that parapsychologists provide strictly repeatable experiments (or strictly replicable findings) before they bother to take a careful look at the evidence. The impossibility of strict repeatability, on the other hand, does not imply that we should stop worrying about repeatability of our experiments or replicability of our findings. Quite to the contrary, there can be no question that, in the long run, we are under the obligation to have our experiments as repeatable as (absolute) 'methodological restrictions' allow. Futhermore, the present discussion should have made plain that the construction of adequate terminological means for a science of parapsychology is by no means a cura posterior (an indefensible opinion, which too many parapsychologists still enjoy). Finally, the considerations and differentiations and terminological suggestions advanced in this paper can also serve, I think, for the formulation of questions regarding our experiments (including our experimenting actions and the ways we talk about them) that are more pertinent than those we have hitherto been trying to answer.

NOTES

1. There were, of course, earlier sporadic and more or less standardized laboratory experiments in parapsychology, such as those by Richet, Coover, Troland, Brugmans/Heymans/Weinberg, Estabrooks, Jephson, and several others. However, standardized experimental techniques became a routine in parapsychological research only after the establishment of the Parapsychological Laboratory at Duke University under Rhine (cf. Rhine, 1934; Mauskopf and McVaugh, 1980). Therefore, in my survey of the discussions about the repeatability problem (Hövelmann, 1983b), I only considered those arguments advanced during the period 1934-1981.

2. An extensive survey and a tentative classification of these various

intermediate positions are to be found in Hovelmann (1983b, esp. pp. 33-43).

3. I had initially used the phrase 'methodically justifiable' as an equivalent of the German phrase 'methodish rechtfertigbar'. However, I have been told that there is a slight difference in meaning between the German word 'methodish' and the English word 'methodical(ly)', and that, because of these different connotations, use of the phrase 'methodically justifiable' may not be entirely appropriate in this context. Therefore, I have decided to substitute 'systematical(ly)' for 'methodical(ly)'. I have already used the phrase 'standardized, methodically constructed terminology' in several other places (for instance, cf. Hövelmann 1983c, in press - a). In order to keep my own use of words as consistent as possible across publications and in order to avoid possible confusion, I have retained the word 'methodical(ly)' in contexts where I am dealing with the construction of adequate terminological means.

4. Some physicists, such as Werner Heisenberg, in particular, understood this question in a literal sense.

5. Translations are mine. Höfling's textbook of physics, which I have quoted several times, is by no means an exception. As Tetens (1977) has exemplarily pointed out, similarly unreflected statements are also to be found in most of the other contemporary textbooks of physics.

6. A discussion of the relationship between experimentation and observation is beyond the scope of the present essay. I am planning to deal with this question in a subsequent paper.

7. For a discussion of the scientific status of contemporary psychology, see Janich (1981).

8. The outcome of an experiment is to be called 'negative' if it does not exactly correspond with the outcome that has been explicitly formulated and predicted in advance.

9. This is completely compatible with the way experimental laboratory research is actually conducted. Suppose that the usage of a measuring tool (of a balance, for instance) in a physical experiment leads to contradictory results. From this the experimenter will, of course, not infer that the 'laws of nature' must have suddenly changed and that the distinction of weights by means of balances will be impossible

henceforth. He will rather assume that the balance is defective and, therefore, replace it by another one that conforms to his norms for tool functioning.

10. The given scheme was first derived and founded by Holm Tetens in his philosophical dissertation on a technically-oriented reconstruction of classical mechanics (Tetens, 1977, esp. chapters 1 and 4). This scheme also covers experiments which do not result in an exactly defined single situation S_2 (e.g., experimental investigations of currents of liquid and gaseous substances or of radio-active decay). In experiments of this type, there merely turn up deficiencies in the procedure of realization: it is impossible in these cases to obtain any particular pre-specified situation S_2 by means of systematic variation of the situation S_1. Therefore, it is necessary here to count out the frequencies of the occurrence of $S_2,\ldots,S_n$ (i.e., $p(S_2),\ldots,p(S_n)$). This results in the following extension of the schematic representation of the standard form of experiments in the experimental sciences:

$$\begin{array}{ccccc} & & & & p(S_2) \\ & & & & \vdots \\ S_0 & a & S_1 & C & \vdots \\ & & & & \vdots \\ & & & & p(S_n) \end{array}$$

11. Janich (1981, pp. 435-436) points out that, with regard to their success, "descriptive and prescriptive versions of recipes are equivalent". He emphasizes, moreover, that his "term 'knowledge about recipes' ... is to express that the linguistic composition of 'know-how' is to make this 'know-how' teachable for the purpose of further application by the instructed person" (my translation).

12. 'Apparatus', derived from the Latin word 'apparatus', literally means 'something that has been prepared'.

13. If I am right with what I have said so far, this may have consequences for various attempts of recent years to provide (quantum-) physical models for parapsychological phenomena. I am planning to discuss this question elsewhere. In any case, it would

cause considerable problems to (meta-) physical theories of the Cappa and Bohm variety.

14. For a terminological distinction between 'action' and behaviour', see for instance Hövelmann (1984, pp. 150-151).

ABSTRACT

Parapsychologists and their critics alike commonly consider the question of repeatability to be one of the most momentous problems facing parapsychology. However, despite the well-recognized importance of this question, discussions of the repeatability problem have mostly been extremely confused, crucial terms have frequently been used in non-uniform and highly ambiguous ways, and both parapsychologists and their critics seem to be content with a diffuse comprehension of the nature and purposes of experimentation. Therefore, an attempt is made (along with a number of terminological suggestions and clarifications) to suggest a conceptual framework for future discussions of the repeatability problem. In the first part of this contribution, a conceptual framework for the discussion of experimentation and of repeatability of experiments for a science like physics is suggested by way of a critique of common realistic and naturalistic misconceptions of the natural sciences. This suggestion rests on the insight that any empirical scientific knowledge decisively depends on the experimenters' active interventions in given situations. In addition, the crucial role language plays in empirical sciences is emphasized. In the second part, it is pointed out that experimentation in sciences like psychology and parapsychology is subject to a number of methodological restrictions. A distinction is made between relative methodological restrictions (which are rooted in insufficiencies of our current practice or theories) and absolute methodological restrictions (which spring from differences between what here has been termed 'experimental' and 'experimenting' scientific activities). The discussion of absolute methodological restrictions shows why strict repeatability cannot be achieved in a science like parapsychology. The hope is expressed that the conceptual framework developed in this paper will contribute to better organized and less confused future discussions about the perennial problem of repeatability of parapsychological experiments.

REFERENCES

Albert, H. 'Traktat über kritische Vernunft', 4th, improved edition, Tübingen, J.C.B. Mohr (Paul Siebeck), 1980.

Beloff, J., Blackmore, S.J., Braud, W.G., Edge, H.L., Honorton, C., Rao, K.R., Schlitz, M., Stanford, R.G. and Walker, E.H. 'The repeatability problem in parapsychology', Parapsychological Review, 1983, 14, (6), 1-4.

Blackmore, S.J. 'Unrepeatability: parapsychology's only finding', Paper presented at the 32nd International Conference of the Parapsychology Foundation, San Antonio, October 1983.

Broughton, R.S. 'Repeatability and experimenter effect: are subjects really necessary?', Parapsychology Review, 1979, 10, (1), 11-14.

Broughton, R.S., Millar, B., Beloff, J., and Wilson, K. 'A PK investigation of the experimenter effect and its psi-based component', in Roll, W.G. (Ed.) 'Research in Parapsychology 1977', Metuchen, N.J. and London, Scarecrow Press, 1978, 41-48.

Dingler, H. 'Die Ergreifung des Wirklichen', Munich, Eidos, 1955.

Hempel, C.G. 'Philosophy of the Natural Sciences', Englewood Cliffs, N.J., Prentice-Hall, 1966.

Höfling, O. 'Lehrbuch der Physik', 9th edition, Bonn, F. Dummler, 1971.

Hövelmann, G.H. 'Book review of "Parapsychology and the Experimental Method" (edited by B. Shapin and L. Coly)', J.o.P., 1983, 47, 68-74. (a)

Hövelmann, G.H. 'Zum Problem der Wiederholbarkeit parapsychologischer Experimente', Zeitschrift f. Parapsychologie u. Grenzgebiete d. Psychologie, 1983, 25, 29-54. (b)

Hövelmann, G.H. 'Seven recommendations for the future practice of parapsychology', Zetetic Scholar, 1983, 11, 128-138. (c)

Hövelmann, G.H. 'Animal Loquens. Kritische Uberlegungen zu Vorannahmen, Methoden und Interpretationen in einem Teilbereich der Vergleichenden Psychologie', Unpublished Master's Thesis (v + 187 pages), Marburg/Lahn, Institute for Philosophy, Philipps University, 1984 (a slightly enlarged English version of this Thesis is scheduled for publication in Semiotica).

Hövelmann, G.H. 'A constructively rational approach to parapsychology and scientific methodology (Responses to my commentators and some further attempts at clarification)', Zetetic Scolar, in press. (a)

Hovelmann, G.H. 'Research and presentation in science: A reply to Dr. Scott's "further comments"', Zetetic Scholar, in press. (b)

Hövelmann, G.H. 'Against historicism: Critical remarks on Thomas Kuhn's conception of science and its reception in parapsychology', J.o.P, in press. (c)

Hövelmann, G.H. 'Technikorientierte versus weltbilderzeugende Naturwissenshaft, Marginalien zu einigen Auffassungen W. Buchels', in press. (d)

Janich, P. 'Die Sprache der Physik und die Wirklichkeit der Naturwissenschaften', Dialectica, 1977, 31, 301-312.

Janich, P. 'Die Protophysik der Zeit. Konstruktive Begrundungen und Geschichte der Zeitmessung', Frankfurt/M., Suhrkamp Verlag, 1980 (English version: Dordrecht, D. Reidel, in press).

Janich, P. 'Ist die Psychologie auf der Grundlage technischer Rationalitat als Wissenshaft moglich?', in Kempf, W. and Aschenbach, G. (Eds.) 'Konflikt und Konfliktbewaltigung', Berne, Stuttgart, and Vienna, Huber, 1981, 419-441.

Lorentz, K. and Mittelstrass, J. 'Die Hintergehbarkeit der Sprache', Kant-Studien, 1967, 58, 187-208.

Mauskopf, S.H. and McVaugh, M.R. 'The Elusive Science. Origins of Experimental Psychical Research', Baltimore and London, The Johns Hopkins University Press, 1980.

Millar, B. 'The distribution of psi', E.J.P., 1979, 3, 78-110.

Neppe, V.M. 'Extrasensory perception: An anachronism and anathema', J. S.P.R., in press.

Palmer, J. 'Critics of parapsychology: Some common elements', Paper presented at the 33rd International Conference of the Parapsychology Foundation, New Orleans, August 1984.

Popper, K.R. 'The Logic of Scientific Discovery', London, Hutchinson, 1959 (first published in German, 1935).

Popper, K.R. 'Die beiden Grundprobleme der Erkenntnistheorie', Tubingen: J.C.B. Mohr (Paul Siebeck), 1979.

Rhine, J.B. 'Extra-sensory Perception', Boston, Boston Society for Psychic Research, 1934.

Shapin, B. and Coly. L. (Eds.) 'Parapsychology and the Experimental Method', New York, Parapsychology Foundation, 1982.

Stokes, D.M. 'Books review of 'Mind at Large' (edited by C.T. Tart, H.E. Puthoff, and R. Targ)', J. A.S.P.R., 1980, 74, 445-454.

Tart, C.T., Puthoff, H.E., and Targ, R. (Eds.) 'Mind at large. (Institute of Electrical and Electronic Engineers Symposia on the Nature of Extrasensory Perception)', New York, Praeger, 1979.

Tetens, H. 'Bewegungsformen und ihre Realisierungen. Wissenschaftstheoretische Untersuchungen zu einer technikorientierten Rekonstruktion der klassischen Mechanik', Philosophical Dissertation (312 pages), Erlangen, Department of Philosophy, Friedrich-Alexander University, 1977.

Tetens, H. 'Was ist ein Naturgesetz?', Zeitschrift fur allgemeine Wissenschaftstheorie, 1982, 13, 70-83.

Thalbourne, M.A. 'The conceptual framework of parapsychology: Time for a reformation', Paper presented at the 27th Annual Convention of the Parapsychological Association, Dallas, August 1984.

Gerd H. Hövelmann,
Rollwiesenweg 42,
3550 Marburg/Lahn,
West Germany

28th ANNUAL CONVENTION OF THE PARAPSYCHOLOGICAL ASSOCIATION

The 28th annual convention of the Parapsychological Association will be held August 12-16, 1985, at Tufts University in Boston, Massachusetts. This convention celebrates the centennial year of the American Society for Psychical Research. Persons interested in attending the convention may write to the convention's Arrangements Chairperson, Fannie Knipe, A.S.P.R., 5 West 73rd Street, New York City, New York 10023, U.S.A..

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

Papers should be equivalent to full-length journal papers and adhere to the style of the 'Publication Manual of the American Psychological Association' (2nd edition). They must be typed, double-spaced, and in camera ready form for inclusion in the Convention proceedings. Four copies of proposed theoretical, methodological or philosophical papers must reach the Program Chairman by April 19, 1985. The deadline for empirical papers is May 24, 1985 to provide authors more time to complete data collection and analysis. Authors should supply an estimate of their presentation time, not to exceed 30 minutes, excluding a question period. Indicate, where applicable, which of the authors will make the presentation. Papers may not be presented in absentia, although in special cases they may be published in the

Convention proceedings.

Posters are brief papers or other materials presented on poster board. Proposals for posters must include four copies of all material to be presented on the poster and the size of the required posters. Photocopies of photographs are acceptable. This material must reach the Program Chairman by May 24, 1985.

Members and associates of the Parapsychological Association may propose a symposium, panel discussion, or workshop. Symposia are formal presentations by participants on related topics. Proposals for symposia must include four copies of a summary sheet indicating the title, chairperson, participants, order of presentation, and proposed time allotments, up to a total of 90 minutes, including question period, and a full paper, prepared according to the preceding directions, from each participant. This complete package must reach the Program Chairman by April 19, 1985.

Panel Discussions are informal discussions designed to maximize spontaneous interactions between the participants and the audience. Formal presentations should be limited to 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. This package must reach the Program Chairman by May 24, 1985.

Workshops are informal presentations, discussions, or demonstrations. Proposals for workshops must include a summary sheet including a title, chairperson, participants, and workshop activity. Workshops will not be listed as part of the formal Convention program. This summary sheet must reach the Program Chairman by May 24, 1985.

Address all correspondence regarding the program to:

Dr. Dean Radin, Chairman, PA Program Committee,
5865 Timber Drive, Columbus, Ohio 43213, U.S.A.

DR. ROBERT L. MORRIS, APPOINTED PROFESSOR OF PARAPSYCHOLOGY AND HOLDER OF THE KOESTLER CHAIR AT THE UNIVERSITY OF EDINBURGH

Just before this issue went into print the important and gratifying news arrived that Dr. Robert L. Morris, 42, had been appointed as professor of parapsychology and holder of the Koestler Chair, established at the University of Edinburgh.

As the readers of the E.J.P. may recall, I stressed in my editorial in the May 1984 issue the great importance for our field of the establishment of the Koestler Chair. It certainly constitutes one of the landmarks in the history of parapsychology. It is also very much gratifying that the pioneering work, sometimes done under rather difficult conditions, which Dr. John Beloff has carried out has obtained due official recognition by the establishment of the Chair in Edinburgh.

Dr. Robert L. Morris has been a highly regarded parapsychologist who has distinguished himself as a very capable and productive research-worker and creative scientist in a wide range of parapsychological areas. A very learned parapsychologist and an excellent lecturer, he has also distinguished himself as a very capable policy-maker in our field by the fine services he has rendered the Parapsychological Association in several of its leading functions. Besides his very broad knowledge of the subject matter one of his great assets as a research leader and educator is his much appreciated good sense of humour. In addition to offering leadership, he is generally recognized as a person who can cooperate very smoothly, in spite of the fact that he is a person of clear-cut opinions which he is willing to defend.

My acquaintance with Robert Morris goes back to a stay at the Parapsychology Laboratory in Durham in 1963. He was then a very young man who came to the lab for a shorter stay. I believe that he had brought a simple PK-machine from Pittsburgh (or was it Philadelphia?) to Durham, an apparatus that utilized marbles instead of dice for his test of PK. Two summers later we both belonged to the same staff at Dr. Rhine's lab, and I am rather sure that Dr. Beloff and Dr. Morris met for the first time that summer. One did not need to possess any precognitive power to predict that Robert Morris would make an illustrious career in parapsychology. His inventiveness and dedication made that conjecture very probable. To people in our field his activities during the last 20 years should be rather well-known, too well-known to warrant much rumination.

I believe that Robert Morris and his wife, Joanna, also well-known in parapsychological circles, should be prepared for some quite noticeable cross-cultural differences. Scotland is rather different from the U.S., and Scotland differs in many ways from continental Europe both as regards geography and mentality. Nevertheless, even if there may exist characteristic differences as regards attitudes towards the paranormal between the U.S. and Europe, especially among the layman, by and large there are more similarities than dissimilarities. In spite of the fact that Europe has been somewhat more permissive as regards university-attached parapsychology than the U.S., he should be quite aware that parapsychology is still a beggar within the academic establishment. There is still no general acceptance of parapsychology as a legitimate scientific endeavour within academia in Europe, especially not among psychologists. However, European universities have shown the degree of open-mindedness that makes university-attached research possible. At best one can say that parapsychology is allowed to be on trial. Any other conception would not only be an overestimation of the actual situation but a most dangerous attitude!

As we all know there exists a historical and psychological connection between parapsychology and certain religious beliefs. The founding fathers of 'psychical research' made no secret of the fact that they hoped to defeat materialism and agnosticism of their time by means of research findings in parapsychology. This emotionally coloured interest in metaphysical issues related to parapsychology may explain much of the rather aggressive and uncompromising attitudes that one finds in the two camps: 'super-sheep' and 'super-goats'. In both the cases they can be characterized as 'fundamentalists'.

After more than eleven years as a chair-holder of parapsychology I do not hesistate to state that the 'super-sheep' have caused me more embarrassment and annoyance than the die-hard critics, in spite of the fact that they often have described themselves as 'friends' and 'supporters'! In a highly critical academic setting like the one that Dr. Morris will face at the University of Edinburgh, I believe that it would be wise to make it quite clear that the 'super-sheep' cannot count on much attention since the most important thing is to get sufficient time for meaningful research. Argumentations with the critic who already knows everything is just as much wasted time and effort! Parapsychology may have some consequences for issues which today certainly are more or less metaphysical in its nature, but issues which by means of conceptual and technical breakthroughs may turn assessible to empirical tests. But whether that is going to happen or not we simply do not know. I am sure that Dr. Morris will be the right man to display good judgment in his dealing with the two types of 'fundamentalists'. Furthermore I am quite sure that he is very much aware that even in the case that it would turn out that most of the effects parapsychologists have reported are methodological artefacts, even such a finding is of great interest and help to science. One important aspect of our research, I believe, is to find out the degree of relevance and the limitations of efficacy of methodology and research strategies utilized (and honoured) in ordinary science as well as in experimental parapsychology. It is also a psychological as well as sociological interest to know more about what makes people attribute to certain experiences a paranormal nature.

I am looking forward very much to a close, fruitful and long-lasting cooperation between the Utrecht laboratory and the Edinburgh research establishment. I also hope that the idea of running a European regional branch of the Parapsychological Association may find a strong supporter in Bob Morris. As a matter of fact the first step towards the creation of the E.R.P.A. was taken several years ago (see E.J.P.,3,1,November 1979), although, for different reasons the project has been shelfed. I have great expectations about the vitalizing effect that the establishment of the Koestler Chair will have not only within the U.K. but for Europe in general.

I take much pleasure in congratulating all the parties concerned as regards the appointment of Dr. Robert L. Morris as the holder of the chair. It goes without saying that Robert Morris and his very nice and knowledgeable wife, Joanna, should feel a very warm welcome to Europe!

Martin Johnson

EUR.J.PARAPSYCHOL., V.5, 1985

EXPERIMENTER EFFECTS IN A PLETHYSMOGRAPHIC ESP EXPERIMENT

Joop M. Houtkooper
Erlendur Haraldsson

INTRODUCTION

Plethysmograph recording of blood volume in the fingers of percipients has been used as an indicator of ESP in previous experiments. For an overview of these experiments up to 1980, see Haraldsson, 1980.

The experimental method followed by the second author in his previous experiments (Haraldsson, 1970, 1972, 1980) is described as follows: The plethysmograph recording was made at a time when an agent in another room was instructed to concentrate on names of emotional significance to the percipient. During an equal number of control trials, the agent had no names on which to concentrate. A pair of subjects participated in each session, one of them first as percipient and the other as agent, immediately after which they changed roles, the first percipient taking the role of agent and the first agent becoming the second percipient.

In a series of these experiments (Haraldsson, 1970, 1972), it was found that the first percipients obtained scores below mean chance

----

This paper fulfils the publication policy of this journal.

expectation (MCE) and that the second percipients scored above MCE, and to such an extent that the scores of the second percipients were significantly higher than those of the first percipients. This differential scoring pattern was termed "percipient-order effect". In the experiment presented in Haraldsson (1980) the percipient-order effect was confirmed, whereas overall scoring did not deviate significantly from zero. However, if overall scoring was analyzed for the 8 different experimenters, there was an indication of systematic differences between experimenters at a significant level. This fortuitous finding reminds of the occurrence of the checker-effect, such as found by Feather and Brier (1968), and of the induced difference in scoring between groups of differently motivated experimenters in an experiment by Taddonio (1976). For reviews of experimenter effects in parapsychological research, see White (1976a, b) and Kennedy and Taddonio (1976).

The findings of the previous experiments with plethysmographic ESP pose the following questions: First, what is the essential condition producing psi-missing in the first percipient of the pair and psi-hitting in the second? Secondly, how does the experimenter-effect come about? Is it produced by the social-psychological interaction? (Moreover, the subjects were recruited from the acquaintances of the individual experimenters.) Or, is it an experimenter psi-effect, as apparent in Feather and Brier (1968) and Taddonio (1976)? Moreover, can we find evidence that these effects are produced by the observation of the outcome by the experimenter, as is the view of observational theory (see Houtkooper, 1983)? These questions are dealt with by the present experiment.

EXPERIMENTAL CONDITIONS

In the previous experiment, it was attempted to elucidate the causes of the percipient-order effect by including two experimental conditions, the length of the period of rest before the start of the sending by the agent, and, whether or not ganzfeld stimulation was given to the percipient. However, these conditions apparently had no effect, whereas the percipient-order effect was significant. Therefore, in the present experiment another condition is tried: a short test of conscious ESP is carried out either before or after the plethysmograph sessions. This conscious ESP-test, of the forced choice type with immediate feedback, is used as an "appetizing" condition for the plethysmogram experiment. It is not planned to analyze the data from this ESP-test itself, since it constitutes only a small number of trials, not warranting the test of any hypothesis on this task. The task is therefore only used as an experimental condition, in order to attempt to manipulate the psi-missing observed previously in the first-percipients, as opposing the psi-hitting observed in the second percipients of the pairs of subjects. Furthermore, the experimenters function as an experimental condition, as will be exposed in the following sections.

SUBJECTS AND EXPERIMENTERS

The 8 experimenters in the present study were students who carried out this experiment in the course of their studies at the University of Iceland. They recruited the 80 subjects from amongst their acquaintances. The experimenters worked in pairs, the assistant experimenter staying with the agent and handing him the targets (names of persons written on cards) at predetermined intervals, while the chief experimenter attended to the percipient and took the plethysmographic measurements. In the analysis of the experiment, the experimenters acted as independent measurers/checkers of the data for subjects other than the ones they handled as chief experimenter.

MEASUREMENTS OF THE PLETHYSMOGRAMS

The scores are derived from the plethysmograms as described previously (Haraldsson, 1980), by measuring the "dip" in the experimental and control periods, while the measurer is blind to the condition. After checking with the target list, experimental scores are compared with control scores by the Wilcoxon matched-pairs signed-ranks test, resulting in a z-score per subject. In the present experiment the plethysmogram of each subject was measured three times, first by the chief experimenter of the session, secondly by one of the other experimenters (called the checker), chosen according to a balanced scheme (see table 1).

This scheme was designed beforehand but unknown to the experimenters at the time of the experimental sessions. Thirdly, as the first two measurements are carried out by student-experimenters, their results

TABLE 1
Allocation to subjects, of experimental condition and of experimenters and checkers, for one group of four experimenters.

CHIEF EXPERIMENTER	SUBJECT NUMBER	ESP-TEST BEFORE/AFTER	PERCIPIENT ORDER	CHECKER
E1	1	A	1	E2
E1	2	A	2	E3
E1	3	B	1	E4
E1	4	B	2	E2
E1	5	A	1	E3
E1	6	A	2	E4
E1	7	B	1	E2
E1	8	B	2	E3
E1	9	A	1	E4
E1	10	A	2	E2
E2	11	B	1	E1
E2	12	B	2	E4
E2	13	A	1	E3
E2	14	A	2	E1
E2	15	B	1	E4
E2	16	B	2	E3
E2	17	A	1	E1
E2	18	A	2	E4
E2	19	B	1	E3
E2	20	B	2	E1
E3	21	A	1	E4
E3	22	A	2	E1
E3	23	B	1	E2
E3	24	B	2	E4
E3	25	A	1	E1
E3	26	A	2	E2
E3	27	B	1	E4
E3	28	B	2	E1
E3	29	A	1	E2
E3	30	A	2	E4

TABLE 1
(continued)

CHIEF EXPERIMENTER	SUBJECT NUMBER	ESP-TEST BEFORE/AFTER	PERCIPIENT ORDER	CHECKER
E4	31	B	1	E3
E4	32	B	2	E2
E4	33	A	1	E1
E4	34	A	2	E3
E4	35	B	1	E2
E4	36	B	2	E1
E4	37	A	1	E3
E4	38	A	2	E2
E4	39	B	1	E1
E4	40	B	2	E3

were checked again by the second author, so that there was the least possibility of measurement errors influencing the result. In no way should the subjective measurement errors be confused with the experimenter effects of interest here. Therefore, the finally rechecked z-scores will be used in the subsequent analysis as the subjects' ESP-scores.

EXPERIMENTER EFFECTS

The general hypothesis is that there will be systematic differences between the scores obtained by different experimenters. Here we assume these differences to be present not only in connection with the chief-experimenters, who dealt most with the percipients and who scored their plethysmograms for the first time, but we also expect differences in scoring with different second measurers/checkers.

About the causes of these experimenter effects, it has to be kept in mind that, the effect of the chief experimenter might be caused by the experimenters' selection of subjects, by their way of handling them, or by the experimenters' psi. In contrast, the effect of the checker can only be caused by the checkers' psi.

In the present experiment we pool the effects of the chief experimenter and the checker to obtain an overall experimenter effect.

HYPOTHESES

H1: The percipient-order effect; to be tested one-tailed in the direction that the second percipient scores better than the first by a matched-pairs t-test on the pairs of subjects, at the .05 level of significance.

H2: The "appetizing" effect of having a short test of conscious ESP before the plethysmogram sessions as compared with having it afterwards; to be tested by an analysis of variance on the z-scores of the first-percipients, the independent variable of interest being "ESP-appetizing". In the analysis of variance the variables "Experimenter" (i.e. chief-experimenter) and "Checker" (i.e. second measurer/checker) are included, as their effects might otherwise enhance the residual (error) variance. The two-way and higher interactions will be pooled with the residual variance. The F-value obtained for the main effect of ESP-appetizing will be tested at the .05 level.

H3: Experimenter effect; to be tested by an analysis of variance on the z-scores of all subjects with the independent variables:

1. Percipient-order.
2. ESP-appetizing.
3. Experimenter.
4. Checker.

As the experimenters are divided into two pools of 4, the analysis of variance is done separately for the two pools. In each, Experimenter and Checker have 3 degrees of freedom. Percipient-order and ESP-appetizing are included to avoid spurious enhancement of the residual variance. Also, their interaction is excluded from the residual variance, because both variables concern hypothesized effects. The other two-way interactions and the higher interactions are included in the residual variance. This gives 30 degrees of freedom for the residual variance per experimenter pool. Finally, the effects of Experimenter and Checker are pooled, and both experimenter pools are pooled, giving 12 degrees of freedom for the overall-experimenter effect and 60 for the residual variance. The resultant F-value is to be tested at the .05 level of significance.

RESULTS

The results for the hypotheses are:

H1: The percipient-order effect.
The t-test results in: t=+0.12; 39 degrees of freedom; n.s.

H2: The "appetizing" effect.
The analysis of variance on the z-scores of the first-percipients results in: F(1,15)=0.110; n.s.

The average z-scores of first-percipients having had the conscious ESP-task before and after the plethysmogram sessions were -0.37 and +0.03 respectively. Thus the, non-significant, effect is opposite the expected direction.

H3: Experimenter effect; we give the analysis of variance in table 2.

The F-test on the overall experimenter effect results in: F(12,60)=1.730; p=.088; n.s.

In table 2, it appears that some of the contributions to the overall-experimenter effect are themselves significant:

1. The overall checker effect: F(6,60)=2.638; p=.025.
This is quite unexpected, since the checkers have only their psi to produce an effect, as compared with the experimenters who can also do this by normal means.

2. The overall experimenter effect in the scores of the second experimenter group is almost significant: F(6,30)=2.378; p=.053.
This effect is again most strongly produced by the checkers in the second experimenter group, where it is significant: F(3,30)=3.706;

TABLE 2
Analysis of variance for the experimenter effects

PERCIPIENT	EXPERIM. GROUP	DEPEND. VARIAB.	INDEPEND. VARIABLE	DF	M.S.	F	P
1 + 2	1 .. 4	Z	EXPTER	3	3777.2	0.551	.651
			CHECKER	3	9274.9	1.353	.276
			E + C	6	6526.1	0.952	.474
			RESID.	30	6854.6		
1 + 2	5 .. 8	Z	EXPTER	3	8643.8	1.049	.385
			CHECKER	3	30545.7	3.706	.022
			E + C	6	19594.7	2.378	.053
			RESID.	30	8243.0		
1 + 2	1 .. 8	Z	EXPTER	6	6210.5	0.823	.557
			CHECKER	6	19910.3	2.638	.025
			E + C	12	13060.4	1.730	.088
			RESID.	60	7548.8		
1	1 .. 4	Z	EXPTER	3	14822.7	2.725	.091
			CHECKER	3	3667.9	0.674	.584
			E + C	6	9245.3	1.700	.204
			RESID.	12	5438.6		
1	5 .. 8	Z	EXPTER	3	4344.8	0.321	.810
			CHECKER	3	17015.3	1.257	.333
			E + C	6	10680.0	0.789	.595
			RESID.	12	13541.1		
1	1 .. 8	Z	EXPTER	6	9583.7	1.010	.442
			CHECKER	6	10341.6	1.090	.396
			E + C	12	9962.7	1.050	.439
			RESID.	24	9489.8		

TABLE 2
(continued)

PERCIPIENT	EXPERIM. GROUP	DEPEND. VARIAB.	INDEPEND. VARIABLE	DF	M.S.	F	P
2	1 .. 4	Z	EXPTER	3	13243.5	3.038	.071
			CHECKER	3	10672.4	2.448	.114
			E + C	6	11957.9	2.743	.065
			RESID.	12	4359.2		
2	5 .. 8	Z	EXPTER	3	8791.6	1.721	.216
			CHECKER	3	16864.7	3.301	.058
			E + C	6	12828.7	2.511	.082
			RESID.	12	5109.6		
2	1 .. 8	Z	EXPTER	6	11017.6	2.327	.065
			CHECKER	6	13768.5	2.908	.028
			E + C	12	12393.1	2.618	.022
			RESID.	24	4734.4		

p=.022.

3. If the first and second percipients are analyzed separately, the overall experimenter effect is largest in the second percipients: $F(12,24)=2.618$; $p=.022$. This is slightly more pronounced in the checkers: $F(6,24)=2.908$; $p=.028$.

DISCUSSION

The findings in the present experiment are somewhat disappointing, as none of the hypotheses put forward could be confirmed. Especially the effect of the "appetizing" ESP-task appears not at all what was expected. This finding all the more begs the question how the percipient-order effect comes about. However, the percipient-order effect occurred in the present experiment only very weakly. It has to be remarked that the pooled result of the percipient-order effect in the series of experiments including Haraldsson, 1972 and 1980 and the present experiment is still significant: $t(112)=+2.41$, $t(36)=+2.26$ and $t(39)=+0.12$ respectively. To pool these results, the t's can be weighted according to their number of degrees of freedom. The result, if normalized, approximates a Student's t distribution: $t(187)=2.91$; $p<.005$, two-tailed.

About the effects in this study, it might first be remarked that the authors hold different opinions on them. The second author was most keen on the study of the percipient-order effect and thought it very unlikely that the effect of the checker could be real. The first author, as opposed to this, found the previous indications of an experimenter effect likely explained by an observational effect, instead of by a social-psychological or a subject-selection effect. Moreover, the view according to observational theory is that the ESP effects have little, or nothing at all, to do with the psi-abilities of the subjects, as they do not receive feedback about their performance. Rather, it is the opinion of the first author that all possible psi-effects in the present experiment must be due to the experimenters and the later observers of the outcomes, with the only exception being subjects acquiring information on the outcomes inadvertently.

At the time the experiment was designed, in the early summer of 1978, it seemed quite speculative indeed, to do work on the checker effect. The checker effect, if it really occurs, implies an effect occurring after the experimental procedure proper is finished. Moreover, within observational theory (Houtkooper, 1983) it implies a so-called "later-observer" effect, or, "future-observer" effect, as it is called by Weiner (1982) in her survey of the literature for these effects. In the present experiment, the first observation of the experimental outcome is performed by the chief-experimenter. As it were, to alleviate the speculative aspect of the checker effect, the effects of the chief-experimenter and the checker were pooled to get the maximum probability to obtain an effect. The drawback of this choice is that the hypothesized effect has to be broken down into its components to obtain a more clear-cut interpretation.

The result of the experiment is somewhat tantalizing: the overall experimenter effect, as it was hypothesized, did not turn out to be significant, but the checker effect, being of interest for observational theory, actually was significant.

We have seen that the checker effect is most pronounced and significant in the second group of four experimenters and in the second percipients. The effects of experimenters may vary within one group of course more than within another. The second percipients otherwise tend to psi-hitting and it might be hypothesized that, therefore, they tend to score more consistently than the first percipients. Looking at the residual variances of both groups in table 2, this turns out to be the case: the residual variance is much higher in the first percipients than it is in the second percipients. This difference can be tested: $F(24,24)=2.004$; $p=.048$.

The checker effect which is found, could be elucidated by looking at the "measuring errors" in the present experiment. After all, the records were all measured three times, and it is interesting to see whether the checker effect is also present in the measuring errors by the checkers. This idea has not been confirmed by the data. Without going into details, the general trend is that the effects are somewhat less in the measurements by the checkers, whereas they tend to be slightly stronger in the measurements made by the chief experimenters. Especially for the checker effect, this is hard to interpret.

Another way of looking at experimenter effects is by examining the differences that might exist in the percipient-order effect between experimenters. The data were examined for such differences, but no consistent effect appears to be present. The same can be said for the differences between experimenters in the "ESP-appetizing" effect; no indication of an effect whatsoever could be found.

About the scheme of repeated observations as employed in the present experiment, it has to be stressed that this is methodologically not yet perfect. There are however practical limitations on what can be done. The fact that the observers are aware of their role as first or second observer of the data, cannot be avoided very easily in the present setup. Another indication of the practical limitations in a study like this is the fact that two of the student-experimenters dropped out in the time-span of about one month between the experimental sessions and the second measurements of the data and had to be replaced. (This fact does not affect the validity of the analysis of the data presented.)

Furthermore, using the scheme of table 1 involves some confounding of the effects of experimenter and checker, which is accounted for in the analyses of variance presented, but which makes it hard to analyze the individual differences between experimenters. The reason for this confounding is that we would not have the same persons put to the task of measuring the same plethysmograms twice.

For guidelines on some methodologically better designs, which make a more adequate characterization of the individual experimenter possible, see Houtkooper (1983), chapter 6. It might be of interest to note that these guidelines were not available at the time the experiment was designed and that, in fact, they were developed with the experience of the present experiment in mind.

With the above reservations kept in mind, we did try to find out whether the same persons obtained scores in the same direction, both as chief-experimenter and as checker. Using adjusted values for individual contributions, we found for the 6 persons with usable data a correlation of +0.44, which is in the expected direction, but far from significant with N=6.

In conclusion, the percipient-order effect, as has been found in experiments with plethysmographic ESP, has not been confirmed in this study. Neither can we say that we have got a better idea about what the cause of the percipient-order effect might be. Further work, for instance on the physiological states of first and second percipients, can be done on this problem.

About the experimenter effects in this study, it can be said that there is the fairly strong suggestion that a checker effect occurred. This means that the present study at least mildly [Bsupports the theoretical framework of observational theory, and within that, those models that allow for an effect of later observers than the first observer of the experimental outcome.

ACKNOWLEDGMENTS

The following persons acted as co-experimenters in this study: Bjarni Ingvarsson, Ludvik Larusson, Saemundur Hafsteinsson, Halldor Bjarnason, Bjarni Jonsson, Finnbogi Thormodsson, Skafti Jonsson, Sigurdur Sigurbjornsson and Gudrun Arnarsdottir. We wish to thank them for their co-operation in an experiment that put an unusual strain on their patience and endurance.

ABSTRACT

In a series of plethysmographic GESP experiments involving emotional targets (names of persons), it has been found that the subject first acting as the percipient of a pair of subjects participating in a session, obtained lower ESP scores than the other subject, who acted as a second percipient after having had the role of an agent for the first percipient. This differential scoring pattern (the percipient-order effect) was confirmed in a later experiment, in which there was also some indication of systematic differences in subjects' psi scores, depending on which of several experimenters had tested the subject.

In the present experiment, using a formalized checking procedure as part of the experimental design, the authors tested systematically the occurrence of the percipient-order effect, the effect of an "appetizing" conscious ESP task, and the combined effect of different experimenters and checkers on the ESP scores. None of the hypotheses was confirmed but one component of the overall experimenter effect, the effect of the checkers, was significantly present in the ESP scores.

REFERENCES

Feather, S.R., Brier, R.M. 'The possible effect of the checker in precognition tests', J.o.P, 1968, 32, 167-175.

Haraldsson, E. 'Psychological variables in a GESP test using plethysmograph recordings', Proceedings of the Parapsychological Association, 1970, 7, 6-7.

Haraldsson, E. 'Vasomotorische Reaktionen als Indikatoren Aussersinnlicher Wahrnehmung', ['Vasomotor reactions as indicators of extrasensory perception'], Doctoral dissertation, Albert-Ludwigs-Universitaet, Freiburg, 1972.

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

Houtkooper, J.M. 'Observational theory: A research programme for paranormal phenomena', Lisse, Holland: Swets & Zeitlinger, 1983.

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

Taddonio, J.L. 'The relationship of experimenter expectancies to performance in ESP tasks', J.o.P, 1976, 40, 107-114.

Weiner, D.H. 'The once and future psi source: A review of the literature on future-observer effects', SERPA Conference, 1982, Abstract in: J.o.P, 1982, 46, 58-59.

White, R.A. 'The influence of persons other than the experimenter on the subject's scores in psi experiments, J. A.S.P.R., 1976, 70, 133-166. (a)

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

Joop M. Houtkooper
Jan Swammerdam Institute
University of Amsterdam
Meibergdreef 15
1105 AZ Amsterdam
The Netherlands

Erlendur Haraldsson
Department of Psychology
University of Iceland
Sudurgata
Reykjavik
Iceland

EUR.J.PARAPSYCHOL., V.5, 1985

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART I. EXPERIMENTAL DESIGN

Hendrik G. Boerenkamp
University of Utrecht

The history of experimental parapsychology is first of all a history of attempts to find reliable and consistent correlates of psi. Psi is the assumed capacity to obtain in an unknown, non-sensorial, way factual information about the outside world. The assumption that most persons have psi abilities was strongly advocated by Rhine (see Nilsson, 1975). Till then it was generally accepted that although ordinary persons might have spontaneous paranormal experiences, the intentional acquisition of paranormal impressions would be a rare and specific capacity of specially gifted people, i.e. psychics and mediums. They alone claim to be able to obtain intentionally paranormal impressions about facts in the lives of persons unknown to them. Rhine strived to verify his challenging assumption of a general human psi ability with the well-known card-guessing experiments (Rhine, 1934). The novelty of Rhine's approach consisted not only in the assumption that unselected subjects could demonstrate psi abilities, but also in his strict adherence to the paradigm of experimental psychology. The card-guessing technique invited other researchers in the field to collect comparable sets of data, also mainly from groups of unselected subjects. One of the main aims of most of these studies was to find variables which have a predictable impact on the psi performance of subjects. Unfortunately, although in the course of time a great variety of variables have been studied, it turned out not to be possible to find any reliable and consistent correlate of psi performance in card-guessing tests (Schouten, 1973). Neither have such correlates been found with other experimental techniques, as can be concluded from Palmer's exhaustive overview (Palmer, 1978, 1982).

When one considers the relative lack of success in experimental research with unselected subjects, it seems worthwhile to investigate the paranormal ability of psychics. Psychics are subjects claiming to possess paranormal abilities or having produced some suggestive evidence of that, independently of the degree of corroboration of such claims. If psychics turn out to possess specific paranormal abilities and if it is known what the optimal conditions are for them to apply these abilities, research with psychics may provide a better experimental method. Since psychics usually express their paranormal impressions in verbal statements by describing facts in the lives of persons unknown to them, a study of the abilities of psychics involves the analysis of statements made by psychics and how these assumed paranormally based statements are affected by different variables. Nevertheless, although some studies with psychics have been published in the past, no studies have been carried out in which statements of psychics in different experimental conditions have been systematically evaluated and compared.

Systematic studies in general with psychics are rare, mainly for two reasons. The first is indicated by Pratt. In the same period as Rhine's successful experiments, Pratt developed a new method to evaluate quantitatively the verbal statements of psychics (Pratt, 1936). In 1969 he wrote, reflecting on the past (Pratt, 1969, p.9,10):

> "Both publications (Rhine 1934 and Pratt 1936) had a great deal in common.<> They both resulted from the new effort being made to bring parapsychology into the mainstream of the movement toward quantitative experimental methods that was taking hold in academic psychology at the time. Rhine's work presented ESP as an ability that is widely distributed in the general population and as one that lends itself to experimental investigation as do other normal abilities. My work showed through actual research that the non-quantitative material of psychics can likewise be investigated by methods that meet the requirements of modern experimental science.<> In spite of the similarities the effects of the two works are marked by contrasts. Rhine's publication stirred up the kind of interest that guaranteed that the issues it raised would not be neglected and forgotten, and most of the developments in

> experimental parapsychology have in one way or another flowed from that publication. My work on the other hand went virtually unnoticed and the problem area with which it was concerned sank steadily into a deeper and deeper state of neglect."

In other words, from that period on the psychic's ability seemed to be considered as part of a general human capacity and no longer a unique ability, deserving special interest. Not surprisingly then, researchers preferred the relatively fast and efficient card-guessing research technique with unselected subjects over the time-consuming analyses of verbal statements provided by psychics. In addition, from that period on 'good' psychics, who were considered hard to find, often were tested with the card-guessing technique as well.

Another important reason for the absence of systematic studies with psychics derives from the problem of evaluating verbal material. Pratt's method of evaluating verbal statements of psychics quantitatively was better suited to test for the paranormality of statements from one particular psychic than it was in evaluating comparable sets of statements obtained under different experimental conditions from a group of psychics. Subsequent methods which were also developed in the tradition of the 'proof-oriented' research approach were therefore equally unsuitable for experimental studies of a comparative nature.

For instance, the Pratt-Birge method, an adaptation of Pratt's method (Pratt-Birge, 1948), requires records of sessions from a particular psychic for different target persons. The essence of the evaluation is to present all statements in random order to all target persons to check for applicability. A paranormal ability might be attributed to the psychic if the target persons rate statements made about them as correct to a significant degree compared to statements made about other target persons. Hence this method requires that all target persons rate all statements made in the study for applicability, a task which even in a relatively small study quickly becomes too cumbersome. Both Schmeidler (1958) and Roll (1962) applied versions of this method. Hettinger (1940,1941) developed a similar matching procedure but based it on a combined set of statements rather than on individual statements. Calling the complete set of statements about one target person a reading, each reading as a whole is presented to all target persons to check for applicability. To a certain extent this procedure simplifies the task of the target persons. Hettinger's work was critically reviewed by Scott (1949).

Parsons (1949) and West (1949) applied a version of this procedure. Timm (1965) developed a statistical procedure to evaluate statements in so-called chair-experiments which had been carried out by Tenhaeff (e.g. 1953) and Bender (e.g. 1957) with a particular psychic.

Although the Pratt-Birge method is regarded as the standard procedure for evaluating verbal material, it is still in need of much further research (Scott, 1972); moreover, for the practical reason discussed above, it is not well suited for use in an investigation in which the statements of a group of psychics in a number of experimental conditions have to be evaluated. As discussed by Boerenkamp and Schouten (1983), the Pratt-Birge method is based on the assumption that all statements of a psychic in a session with psychic and sitter might be influenced by psi. The attribution of a paranormal character to statements from psychics, provided the psychic has no sensorially acquired knowledge about the target person, depends on two factors. One is the a priori probability of the statement being correct, and the second is the actual correctness of the statement. The Pratt-Birge method and subsequent methods all start with all sitters checking the correctness of all statements. As observed above, this property makes them rather unsuitable for applying in a comparative study.

For the present investigation, a different procedure for evaluating sets of verbal statements was developed (Boerenkamp and Schouten, 1983; Boerenkamp 1984). The procedure is based on a so-called window model. In this model it is assumed that within a session psychics are only occasionally able to apply their psi-ability and that several statements will be based on familiar psychological functions, for instance, the ability to make logical inferences. From sessions reported in the literature one may expect a session to be made up of a substantial number of statements of a general nature (the noise) and a smaller number of specific statements which may be based on psi. These specific statements are assigned potential paranormal value. The term 'potential' is used to indicate that these statements have passed the first requirement, i.e. the a priori probability of the statement being correct is sufficiently low so as to warrant further interest. In this phase of the evaluation it is still not known whether these statements are correct, i.e. whether they do apply to the target person. The procedure requires that for all statements the degree of potential paranormal value is first established by utilizing judges, not target persons. Only statements with a sufficiently high degree of potential paranormal value are retained to be checked for applicability to the target person and thus for further analysis.

It was found that this procedure is feasible because the estimates of degree of potential paranormal value assigned by judges appear to be sufficiently reliable. Because most statements with potential paranormal value, i.e. the statements of low probability, concern specific, observable facts, these retained statements can be checked directly as being true or untrue for the target person for whom they are intended. This procedure allows for the paranormal ability of a group of psychics in a particular experimental condition to be expressed both in terms of the number of statements with potential paranormal value relative to the number of statements of no potential paranormal value as well as in terms of the proportion of statements with potential paranormal value which turn out to be correct.

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. Emotionally, perhaps the most important of such problems involve missing relatives or acquaintances. However, psychics are approached for all kinds of problems in life and often act as consultants for these problems. To help obtain paranormal impressions psychics frequently use objects related to the persons about whom they are consulted. Such an object, for instance a photograph or a ring of the target person, is called an inductor.

Reports of sessions with psychics in the parapsychological literature (e.g. Tenhaeff, 1960, 1962, 1965, 1980) are suggestive of psychics having specific paranormal abilities. However, these reports are not entirely convincing. In the first place, it might be assumed that to a certain extent the sessions have been selectively reported. A fair estimate of a psychic's ability seems possible only if the number of sessions to be held is decided upon beforehand and all sessions are included in the evaluation. However, from the literature it appears that sessions which were considered unsuccessful have most of the time been considered as not of sufficient interest to warrant publication. In the second place, most publications concern the abilities of individual psychics, often the more prominent ones. We do not know how representative these psychics are of psychics as a whole. Thirdly, given the 'proof-oriented' research approach used in the reported studies, most sessions were held under the same conditions. Often different possible sources of information were eliminated in order to be sure that significant results could be attributed to psi.

Usually this was done by presenting the psychic with a sealed envelope containing some object or photograph belonging to the target person and asking for impressions while feedback was not given until after the session. However, these conditions are fairly different from the conditions psychics are used to in their daily practice. For example, in daily life psychics normally receive immediate feedback from the sitter. Since we do not know whether immediate feedback or other variables have a facilitating or perhaps inhibiting impact on the assumed ability, it is difficult to judge whether psychics' activities in daily life are more or less impressive compared to the results reported in the literature. Since it might be assumed that in their normal practice psychics will create optimal conditions for exerting their ability, these conditions were taken as the starting point for this investigation. Both the aims and the design of the investigation were planned and stated before the start of the investigation.

AIMS OF THE INVESTIGATION

The specific aims of this series of studies were:

1) To investigate, based on a predetermined number of sessions and on all statements made in these sessions, whether a sufficiently high number of correct potential paranormal statements are obtained to justify the conclusion that research with psychics offers a more promising research method than the methods currently employed most often in parapsychology.

2) 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.

3) To study in a systematic way the effects of a number of variables listed later in this paper on the statements of psychics and on the number of potential paranormal statements.

4) To compare the number of potential paranormal statements provided by psychics under two different conditions with the number of potential paranormal statements acquired from non-psychics under the same conditions, described later in this paper.

DESIGN

The group of psychics

The group of psychics employed in this study were as representative as possible of all psychics. However, for various reasons it was rather difficult to select a really representative sample. The most important reason was that formal criteria for defining the population of psychics (for instance, a degree) do not exist. Therefore the selection of the group was based on the following considerations: (a) The persons considered themselves psychics and had a wide range of experience in functioning as psychics in sessions with clients. (b) The psychics had to be willing to participate in the studies on terms stated by the researchers. These terms mainly included that they must agree to sessions involving target persons from the environment of the researchers, that a complete tape-recording of all sessions could be made, that they must be willing to participate for a long period of time involving several sessions, and that they must be willing to participate for a relatively small amount of money. (c) Since the number of active psychics was limited, a group of 10 to 15 was considered sufficiently large to warrant conclusions about the ability of psychics in general. One can distinguish between two groups of people who fit the definition of a psychic. On the one hand persons who believe and declare that psi information is received from the deceased (mediums); on the other hand persons who do not make this explicit declaration. The group of psychics selected for this study included people from both categories.

The sessions

To minimize the psychics' feeling that they were taking part in an experimental investigation the following measures were adopted: (a) All sessions were held in the psychic's own home. (b) Sessions were spread out over about five years. Each psychic was called upon about twice a year. (c) The target persons were selected from the environment of the researchers (relatives and friends), just as would be the case normally when clients turn to psychics because of persons in their environment and not merely arbitrarily chosen persons to whom the researchers would have no other interest save their role in the study. Because two experimenters participated, half of the target persons came from the environment of one researcher, the other half from the environment of the other. Accordingly, in each session one researcher acted as sitter and one as the observer. (d) The researchers portrayed themselves as interested but unfamiliar with the way psychics work and adopted a 'pupil's' attitude. Sessions were not introduced to test the psychic's ability but to show the researchers how psychics do their job. (e) Psychics were not informed of the fact that an experiment was being carried out with a predetermined number of psychics and sessions. They were told, however, that now and then some other psychics were visited as well. (f) Each session was presented as a solitary case. Each visit to the psychic started with some social talk, followed by general questions about his or her ability. Then the conversation was directed to particular circumstances the psychic encountered in his or her practice, circumstances which resembled most the conditions for the planned session or sessions for that day. Subsequently the psychic was invited to give paranormal impressions about a target person.

The standard series

Standard series were made up of sessions, one from each psychic, conducted under conditions which resembled as much as possible the conditions psychics normally work in. Thus in view of the second aim of this investigation the data from these standard series could be used to give a description of what usually happens in sessions when clients consult psychics. In order to study the amount of variation in sessions, three standard series were run. The first two standard series, A and B, were obtained at the start of the investigation. Comparing the results of series A and B provided an indication of the typical characteristics of and short-term variation in the behaviour of the different psychics. In addition, the data of these series served as a standard with which to compare the results of the various experimental series. The third standard series, series C, was obtained at the end of the investigation in order to study possible long-term variation in the behaviour of psychics by comparing it with the data from series A and B.

Conditions of the standard series

Describing the conditions of the standard series in more detail is important because some of these conditions were the ones which were systematically manipulated in the experimental series. The following characteristics of sessions psychics hold with clients were considered:

(A) People turn to a psychic either because of problems they have themselves had or because of problems related to someone in their environment. The experimenters did not consult the psychics about themselves, but only about persons from the experimenters' environment. In the standard series the target persons were not physically present in the sessions, thus allowing better control.

(B) People turn to psychics for problems concerning target persons of different age and sex. This might be a relevant variable, for instance, because psychics find it perhaps easier to make statements of potential paranormal value about target persons of their own sex and of their own age. In the standard series a possible systematic influence of these variables was neutralized by utilizing six target persons of different age and sex. The six target persons were men and women of about 25, 45 and 65 years of age respectively. The psychics were split into three equally large subgroups. The target persons were allocated in a systematic way to the psychics of each subgroup. In standard series A and B each psychic held two sessions about two different target persons. In series A the first half of each subgroup of psychics gave a session about the first target person and the second half about the second target person. In series B the first half of each subgroup of psychics gave a session about the second target person and the second half about the first. In addition, the target persons were selected in such a way that the type of relation between researcher and target person was different for each target person. The chosen target persons were a friend, acquaintance, brother-in-law, colleague, uncle and mother-in-law of the researcher, and these six persons were employed in experimental series 1-4 as well. Table 1 presents the allocation of target persons to psychics in the standard series and experimental series.

(C) People sometimes approach psychics for advice about important problems. Indeed, one of the most important problems concerns the unexpected disappearance of a family member or acquaintance. Nevertheless, the majority of psychic consultations concern persons with various problems of a less dramatic character. Therefore, the standard series involved target persons with problems of less acute

TABLE 1
Distribution of target persons over psychics
in standard series and experimental series

Subgroups		Psychics A	Psychics B	Psychics C
Standard series A and B		T 1, M, 25 T 4, F, 45	T 2, F, 25 T 3, M, 45	T 6, F, 65 T 5, M, 65
Experimental series	1	T 6, F, 65	T 1, M, 25	T 3, M, 45
	2	T 6, F, 65	T 1, M, 25	T 3, M, 45
	3	T 5, M, 65	T 4, F, 45	T 2, F, 25
	4	T 2, F, 25	T 5, M, 65	T 4, F, 45
	5	T 7, M, 35	T 7, M, 35	T 7, M, 35
	6	T 8, F, 55	T 8, F, 55	T 8, F, 55
	7	T 8, F, 55	T 8, F, 55	T 8, F, 55
	8	(T 9, F, 35)	(T 9, F, 35)	(T 9, F, 35)
	9	(T 9, F, 35)	(T 9, F, 35)	(T 9, F, 35)
	10	T 10,F, 25	T 11,M, 45	T 12,F, 65
Standard series	C	T 3, M, 45	T 6, F, 65	T 1, M, 25

Notes: T: target person; M: male; F: female.

importance.
(D) People consulting psychics either state their problems to the psychic at the start of the session or expect that the psychic will bring it up in the course of the session based on his assumed paranormal ability. In the standard series the psychic was invited to give his impressions about the target person and no specific problem of the target person was mentioned beforehand. If the psychic asked whether the target person had asked for impressions about certain specific topics the sitter answered "No, not a special topic". Although this deviates from normal practice in consultations it was considered necessary because only in this way could it be ensured that the psychic selected the topics about which paranormal impressions

were acquired. Considering the large differences in age and sex between target persons one can expect large differences between them as regards physical and psychological well-being, civil status, and general circumstances in work and life. Consequently inviting special problem areas from the target persons might have resulted in a variety of different topics, rendering a comparison between psychics and sessions more difficult.

(E) When people consult psychics about someone in their environment the psychic usually requests a photograph of that person or an object, like a ring or pipe belonging to the person, to be used as an inductor. In the standard series the psychic was presented with both a photograph of and an object belonging to the target person. The same photograph of and same object from each target person in the standard series was also used in the experimental series 1-4.

(F) Sitters vary in the extent they give verbal feedback. In the standard series the sitters provided the psychics with feedback because that is what normally happens when clients consult psychics. The psychic first received the information that the sitter was rather well acquainted with the target person. Subsequently the sitter reacted to the statements or questions with feedback. Feedback, or 'informative actions' by the sitter, was given in the form of 'confined informative actions', as well as in the form of 'extended informative actions' in the standard series. 'Confined informative actions' means that the sitter only affirmed or denied the statements with "yes", "no", or "don't know". 'Extended informative actions' means that the affirmation or denial was followed by some clarification. As a rule in this investigation, when giving feedback the sitter avoided as much as possible providing the psychic with information which was not directly related to the topic under consideration. It was not determined beforehand on which occasions the sitter would react with 'confined' or with 'extended' reactions within a session. The sitter would balance as much as possible the number of 'confined' and number of 'extended' reactions.

(G) As is normal in daily practice, sitters may react in a non-verbal way to statements of the psychic. However, the role of non-verbal feedback was not studied in this investigation. Such an investigation, which undoubtedly would be very complicated, can be postponed until it is demonstrated that psychics provide significantly more statements with potential paranormal value in feedback conditions then they do in non-feedback conditions. However, sitters always tried to react verbally to ensure that all feedback was tape-recorded. The researcher who acted as the observer in a session was especially responsible for enforcing this rule because it is to be expected that the sitter may become actively involved in the interaction. Therefore in any case in which the sitter restricted himself to a non-verbal response, for instance by nodding his head, the observer actively interrupted by saying "Oh, is that true?" or "Oh, isn't that true?".

To sum up, in the standard series the psychic was presented with a photograph and object of a target person chosen from the environment of the sitter. The psychic was then invited to give paranormal impressions concerning this person knowing that the sitter was acquainted with the target person. He or she was not informed about special problems in the life of the person. The psychic received immediate feedback to his statements in the form of affirmation or denial, occasionally followed by some clarification, which provided additional related information; however no information unrelated to the topic being considered was provided in the feedback.

The experimental series

The third aim of the investigation was to study in a systematic way the effect of certain variables on the number of correct potential paranormal statements of the psychics. Ten experimental series were carried out in two subgroups of five series each, the second being introduced to the psychics as more 'experimental'. The second subgroup of series was carried out when the sessions of the first subgroup of series were completed. In order to prevent order-effects as much as possible, the administration of each series within each subgroup of series was randomized.

The first five series studied the effect of the following variables:

(1) The type of inductor in feedback conditions (series 1 and 2);
(2) The amount of feedback (series 3 and 4);
(3) The importance of the problem (series 5).

The last five series studied the effect of:

(4) The type of inductor in non-feedback conditions (series 6 and 7);
(5) The non-existence of a target person (series 7 and 8);
(6) The social context (series 8 and 9);
(7) The presence of the target person (series 10).

(1) The inductor in feedback conditions

In series 1, the psychic was given an object belonging to the target person, i.e. a ring or a pipe, to be used as an inductor. In series 2 the inductor was a photograph of the same target person (see also table 1). The psychic was not informed that in both series the same person served as target person. All other conditions were similar to those in the standard series. Since in the two sessions the same target person was involved for each psychic a session of series 1 always preceded a session of series 2.

(2) The amount of feedback

Three different types of feedback or informative actions were distinguished.
a)'No informative actions' meant that the sitter stated beforehand that he knew nothing about the target person and consequently that he was unable to give any feedback. The sitter reacted to statements of the psychic only with "hmm-hmm" and "I see" in order to make clear that he was listening to the statements.
b)'Confined informative actions' meant that the sitter only reacted to statements of the psychic with "yes", "no", or "don't know" (or equivalent reactions). For example:

Psychic: Is something wrong with his blood (?)
Sitter : No, not that I know..

c)'Extended informative actions' meant that the sitter not only affirmed or denied statements but, as clients often do, he expanded on that topic by clarifying the "yes", "no", or "don't know" answer. For example:

Psychic: Is something wrong with his blood (?)
Sitter : Yes, he has diabetes..

In the standard series the sitter reacted with extended as well as confined informative actions. In series 3 the psychic was informed that the sitter was acquainted with the target person. However, the sitter reacted with confined informative actions only. In series 4 the psychic was informed that the sitter was unacquainted with the target person and the sitter reacted with no informative actions only. Target persons for this condition were the persons from the environment of the experimenter who acted in these sessions as observer, but the sitter did not know these target persons. The psychic was not informed that the observer did know the target persons. All other conditions in these series were similar to those in the standard series.

(3) The importance of the problem

Series 5 was planned beforehand in so far as the psychics would be consulted about a very important problem. It was carried out when one of the experimenters became involved in a case concerning a friend of his who had disappeared. Missing persons are one of the most important problems psychics are consulted for. This series was administered at the moment that the occasion arose. For obvious reasons only one target person was involved in this series. These sessions were similar to those of the sessions in the standard series save for the fact that this was a serious case in which one of the experimenters was involved. The psychic was informed at the beginning of the session that the target person was a friend of the sitter and that this friend had disappeared. Then the psychic was invited to provide as much information as possible about the present circumstances of the person as well as to give other impressions concerning the target person.

In contrast to the first five series, the series 6,7,8 and 9 were presented and administered as a related set of sessions comprising an 'experiment' by the researchers. The sessions in these four series differed most from those of the standard series as far as the information available to the psychic was concerned. In none of these four series was feedback given. In each of the series a necklace was used as an inductor. 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 owners of the necklaces. Two 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.

(4) The inductor in non-feedback conditions

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.

(5) The existence of a target person

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 party 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). Paranormal impressions of the psychic in these series could not concern a particular person except perhaps the buyer. In series 8 the necklace was presented without any feedback as the necklaces in series 6 and 7. 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 psychics' impressions.

(6) The social context

In series 9 the psychic was requested to give his or her impressions in the absence of the researchers 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 made by psychics.

(7) The presence of the target person in the session

In series 10 the effect of the presence of the target person at the session was investigated. In this series three additional target persons, aged about 25, 45 and 65 were employed (see table 1). Each target person 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 to bring along a personal object (a photograph and ring) in case the psychic requested one to use as an inductor. If the psychic asked whether the target person had a special problem they were instructed to state: "Not a special problem in particular". They were also requested to avoid providing new information and to limit themselves to feedback about the topic discussed.

Non-psychics acting as psychics

The fourth aim of the investigation was 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 10 to 15 persons matching the group of psychics in age, sex and level of education (Control group 1). 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 10 to 15 persons (doctors, psychologists and laywers) who, because of their profession, were experienced in dealing with problems of people (Control group 2). 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). 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 impressions of the person to whom the object belongs and who is pictured in the photograph. 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 giving impressions about people unknown to them using only a photograph or an object. 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.

ANALYSIS OF A SESSION

The analysis of a session involved three aspects: the informational, structural, and interactional aspects.

The informational aspect

During the sessions psychic and sitter exchanged information about the target person or about someone related to the target person. The information provided by the psychic was assumed to be based on paranormal impressions, i.e. this information was actually a verbalisation of internal sensations. However, even before the session started the psychic had at least some information about the target person including age, sex, and appearance, either because a photograph was provided as an inductor or because the target person was physically present. Then during the feedback sessions more information was added. A typical pattern consisted of the psychic making one or a few statements expressing something about the target person, followed by a statement which invited feedback. For example:

He is an easily distressed person.
He is in good physical health (?)

The sitter usually reacted in such a situation with either confined informative actions (yes,no,don't know) or with extended informative actions in which some information about the topic discussed was added. For example:

Psychic: He has already been missing for several days (?)
Sitter: Yes, eight days ago he left home

Hence, in the course of the session the psychic obtained a growing body of knowledge about the target person. Statements were judged on potential paranormal value by taking into account all the information available at the time of the statement (see Boerenkamp, 1984). A statement was assigned potential paranormal value only if the content was considered as sufficiently specific and spontaneous in view of the available information at the moment the statement was made.

The informational analysis was based on the number of statements with potential paranormal value. Only statements with potential paranormal value were checked on whether they applied or not to the target person. Those which turned out to be correct were called 'statements with positive paranormal value'. Statements of potential paranormal value which appeared to be incorrect were called 'statements with negative paranormal value'.

The structural aspect

The structural aspect related to the various characteristics of the set of statements of a session. These characteristics involved (1) the topics discussed in the statements by the psychic; (2) the number of topics concerning the target person and the number of topics concerning persons related to the target person (e.g. wife, father, child); (3) the percentage of statements related to the past, present, or future; (4) the percentage of statements describing a favourable, neutral, or unfavourable state of affairs; (5) the percentage of statements containing advice; (6) the effects of periods of silence preceding statements; (7) the percentage of positive and rhetorical statements made by the psychic; (8) the number of statements preceding an informative action by the sitter.

The interactional aspect

During the feedback sessions the psychic and the sitter were involved in verbal interaction. In this interaction psychic and sitter were exchanging implicit information about their mutual expectations of the session, these expectations being primarily defined by the complementary type of relation between psychic and sitter. Because of the assumed special ability of the psychic the relationship was of an up-down type. It was the psychic who took the initiative by making statements of a certain content while the sitter was obliged to conform to the psychic's verbal actions. The interactional behaviour of a psychic might be interpreted as a continuous attempt to convince the sitter of the validity of the psychic's ability. The latter will accept this when the statements of the psychic about the target person are specific and correct or probably correct. However, each incorrect statement contradicts the assumed ability of the psychic and thus tends to threaten the basis of the cooperation. Therefore it is to be expected that psychics will take some special actions each time the sitter denies the correctness of a statement. For example:

Psychic: He has already been missing for some days (?)
Sitter: No, it is a week already
Psychic: No?..but that is what I said..some days..not weeks
Sitter: Oh yes, I see..

In this example the psychic resists the denial by suggesting that his statement implies the same information as the feedback of the sitter. The interactional analysis studies the different types of actions the psychic takes after the denial of a statement.

DETAILS OF THE INVESTIGATION

In order to select a group of about 10 to 15 psychics, 24 psychics were approached. Those who believed in an explicit spiritistic explanation, the mediums, were selected from among the best known mediums of the Society of Dutch Spiritists. Originally it was planned to select the subgroup of psychics who do not give an explicit spiritistic explanation for their capacity from psychics who were the best known according to two Dutch parapsychologists, Tenhaeff and Zorab, and from psychics who publicly advertised their claims.

However, it proved impossible to include the latter category in the investigation. Eight persons who publicly advertised their claims were approached but for different reasons none could be included. Four of them turned out to be primarily active in areas which were considered not relevant to this study (for instance, paranormal healing, card reading). Of the remaining four one asked too much compensation, one appeared not to be reliable on the first contact (being absent on the first appointed meeting), one did not want to participate with the excuse of being too busy, and one was not interested because research could lead to nothing good since he viewed his capacity as a gift of God.

Of the best known psychics according to Tenhaeff and Zorab two initially stated their willingness to participate but then withdrew when the project started. They only allowed the researchers to be present and to observe them in their daily practice. Two additional psychics appeared to be especially active as paranormal healers and to restrict their psychic activities to cases of missing persons. Hence it was decided beforehand to approach them only for participation in series 5. Five other well-known psychics were willing to participate. Consequently the group of gifted subjects participating in this investigation consisted of seven mediums, five psychics and in addition two psychics who only participated in series 5. This group consisted of 6 males and 8 females. They ranged in age from 30 to 75 with a mean of about 55 years.

The standard series A and B were carried out with 12 psychics. The experimental series 1,2,3,4 were conducted with 11 of the 12. In series 1 and 2 the contribution of one psychic is lacking because he died before the second session was held. In series 3 and 4 the contribution of one psychic was lost because of a malfunctioning of the tape-recorder in one of the sessions. Series 5 was carried out including the two psychics who were specialized in cases of missing persons. Experimental series 6,7,8 and 9 were collected from 9 psychics. Series 10 and standard series C were carried out with 8. As observed above, one of the psychics died before the start of the second subgroup of series. After a session of series 5 a conflict arose between one of the psychics and the sitter about the compensation for that particular session. The psychic wanted the sitter to approach the wife of the missing person for additional compensation. This claim was refused by the sitter, who was only willing to pay the usual fee. Because of this conflict the psychic refused further cooperation except for experimental series 10. A third psychic ended all activities as a psychic after the death of her only sister. Table 2 presents the participation of the psychics in all series.

TABLE 2
Participation of the psychics in the series

Psychics	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
Series A and B	+	+	+	+	+	+	+	+	+	+	+	+		
Series 1	+	+	+	+	+	+	+		+	+	+	+		
2	+	+	+	+	+	+	+		+	+	+	+		
3	+	+	+	+	+		+	+	+	+	+	+		
4	+	+	+	+	+		+	+	+	+	+	+		
5	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	+	+		+	+	+	+		+	+	+			
7	+	+		+	+	+	+		+	+	+			
8	+	+		+	+	+	+		+	+	+			
9	+	+		+	+	+	+		+	+	+			
10	+	+	+	+	+	+	+		+					
Series C	+	+		+	+	+	+			+	+			

Note: +: participation.

The sessions with the two control groups, each consisting of a group of 12 non-psychics, were carried out by two assistant-researchers.

In addition to the designated series a post-hoc series was carried out with 8 psychics. After the Parapsychology Laboratory was approached in another case of a missing person the occasion was used to call the psychics for a consultation in order to study how telephone sessions differ from face-to-face sessions in the case of a comparably serious problem. In the telephone session the psychics were first asked permission for a tape-recorded registration.

All sessions of the investigation were attended by the author either as the sitter or as the observer, and all recordings were transcribed by the author. Initially it was planned that the transcription of the recordings of the sessions would be carried out by assistants; however, for various reasons this appeared to result in too many inaccuracies in the transcripts. The author made a 'book' for each psychic in which all transcripts of the sessions were ordered in temporal sequence.

The first part of the informational analysis of each session, which consisted of the itemization 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 eight judges was employed. All judges, averaging 23 years of age, were psychology students who volunteered to take part in classes on the evaluation of verbal material. The number of male and female judges involved in the whole investigation was about equal. The itemization procedure as well as the procedure for establishing the degree of potential paranormal value of the statements is described by Boerenkamp (1984).

The second part of the informational analysis, which consisted of the estimation of the positive or negative paranormal value of statements with potential paranormal value, was carried out by the two experimenters. All further analyses were carried out by the author.

ABSTRACT

Experimental research in parapsychology has been rather unsuccessful in finding variables which have a consistent effect on the assumed psi ability of subjects. Based on some basic assumptions made by the 'Rhinean school', research has been mainly carried out with unselected subjects. On the other hand psychics are persons who claim to possess special paranormal abilities which they normally apply to obtain paranormal impressions about persons unknown to them. Hence it might be that experimental research with psychics offers a more fruitful approach than research with unselected subjects.

Systematic and comparative studies of psychics' activities investigated under different conditions appear not to exist. Most studies have concentrated on the ability of one specific psychic and were carried out in the tradition of the 'proof-oriented' approach with the aim to 'prove' or 'disprove' the assumed paranormal ability. The aims of the present series of investigations were more pragmatic. They are a) to study whether the paranormal abilities of psychics

justify the conclusion that research with psychics is more promising than the current research approach, b) to provide a description of the content of sessions when psychics are consulted about a person unknown to them, c) to study in a systematic way the effect of certain variables on the psychic's paranormal impressions, and d) to compare the impressions of psychics with impressions from non-psychics obtained under comparable conditions.

In this paper the design of the entire study is presented. The number of participating psychics and the number of sessions were established in advance. The study involved 3 series of sessions with psychics under conditions which were as similar as possible to the conditions of the normal daily practice of psychics (Standard series). In addition, 10 series of sessions were held in which the role of different variables was investigated. These variables were: the type of inductor, amount of feedback, the importance of the problem, the presence or absence of the target person, and the social context in which the session was held.

EUR.J.PARAPSYCHOL., V.5, 1985

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART II. THE STANDARD SERIES

H.G. Boerenkamp
University of Utrecht

In this paper the results of the standard series A and B are discussed. Among other reasons, the standard series were carried out to provide a description of the content of sessions in which psychics were consulted about persons unknown to them. The sessions are described in terms of number of statements and informative actions (feedback by the sitter) and in terms of the results from the informational, structural and interactional analyses. The results of standard series C will be discussed elsewhere. Series C was carried out at the end of the investigation in order to study possible effects of time on the behavior of the psychics.

Number of statements and informative actions

The length of a session is expressed as the number of statements the psychic makes and as the number of informative actions of the sitter. Informative actions denote statements made and answers given by the sitter as reactions to statements made by the psychic. The total number of statements made by the 12 psychics in standard series A was 1075 and in standard series B 1046. The number of statements by each psychic ranged from 42 to 171 with a mean of about 90 in series A, and ranged from 39 to 128 with a mean of about 87 in series B. The total number of informative actions in series A was 717 and in series B 724. The number of informative actions by the sitter in each session ranged from 21 to 104 with a mean of about 60 in series A, and ranged from 31 to 96 with a mean of about 60 in series B. Hence, we can conclude that series A and B were nearly equal as far as the average number of statements and informative actions is concerned.

The informative actions consisted of 266 extended reactions (37%) and 451 confined ones in series A, and of 302 extended reactions (42%) and 422 confined ones in series B. In extended reactions the sitter reacts to a statement by the psychic with "yes", "no", or "don't know", and by expanding on that topic by clarifying the answer, whereas in confined reactions the sitter limits his reaction to only "yes", "no", or "don't know". The difference in relative number of extended reactions between the two series is not significant (chi-square= 3.02, df=1, n.s., two-tailed). (All chi-square tests applied in this paper are two-tailed).

As expected, the number of informative actions made by the sitter in a session depended on the number of statements made by the psychic. In the 24 sessions of series A and B the sitter reacted with more informative actions when the psychic made more statements (Spearman $r = .71$, $t = 4.83$, $df = 22$, $p < .0001$, two-tailed). (All t-tests applied in this paper are two-tailed).

It also appears that psychics were consistent in the number of statements they made in the sessions. The correlation between the length of sessions of the same psychics in the two series is positive to a significant degree (Spearman $r = .86$, $t = 5.29$, $df = 10$, $p < .001$).

The informational analysis

The procedure of selecting statements with potential paranormal value acquired under feedback conditions 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. Therefore each statement was rated by judges in two different ways on a four-point scale. On these scales the lowest degree of potential paranormal value is represented by a score of 1 and the highest degree of potential paranormal value is represented by a score of 4. In the study mentioned above (Boerenkamp, 1984) it appeared that the distribution of the statements in different categories of potential paranormal value based on the estimates of eight judges did not significantly differ from the distribution obtained when employing only two judges. It was also found that the scores of two judges for the statements correlated significantly with the scores of two other judges. In order to establish a standard for later administration of the procedure with only two judges, the relationship between the scores of two judges was also investigated in that study. It turned out that for 82% of the statements the two judges agreed in assigning either a score of 4 or 3 or a score of 2 or 1. Therefore it was concluded that two judges yield sufficiently reliable estimates of the potential paranormal values of the statements. It also appeared that the judges' ability to estimate potential paranormal value primarily consists of discriminating between specific versus vague statements and between spontaneous versus inferred statements. Apparently it is more difficult to discriminate reliably within categories: between very specific and rather specific statements and between very spontaneous and rather spontaneous statements. It appears to be most difficult to discriminate between very vague and rather vague statements and between very inferred versus rather inferred statements.

The correlation between the scores of the two judges in the standard series appears comparable to the results obtained in the previous study. In series A the judges agreed on 84% of the statements to which of the two categories it should be assigned, in series B the judges agreed on 85% of the statements.

The distribution 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 scales ranging from 1 to 4 are presented in table 1.

According to the Kolmogorov-Smirnov two-sample test the distributions of series A and B are not significantly different (Dmax= .017, Dcrit.01, two-tailed= .071). Applying a cut-off criterion between the medium and medium-high categories, the two standard series yielded a total of 188 statements (9%) of potential paranormal value. After splitting up each session in the first half (of the total number of statements by the psychic) and the second half, it appears that 105 of these 188 statements were made in the first half and 83 were made in the second half of the sessions. This difference is not significant

TABLE 1
Distribution of scores of potential paranormal value in the standard series A and B

score	low 4-7	low-med 7-9	medium 9-11	med-high 11-13	high 13-16	total
series A	586	249	145	64	31	1075
series B	588	230	135	55	38	1046
series A %	55%	23%	13%	6%	3%	
series B %	56%	22%	13%	5%	4%	
total A+B	1174	479	280	119	69	2121
total A+B %	55%	23%	13%	6%	3%	100%

(chi-square= 2.55, df= 1, n.s.).

It appears further that each psychic was rather consistent in the percentage of statements with potential paranormal value (the medium-high and high categories of table 1) they made in the two series. The correlation between the percentage of statements with potential paranormal value over psychics in series A and series B is positive to a significant degree (Spearman r= .76, t= 3.70, df= 10, p< .01). The distributions of all statements (all) and the statements with potential paranormal value (ppv) over psychics are presented in table 2.

The second step in the informational analysis was to establish how many of the 9% statements having potential paranormal value fulfill the criterion of 'sufficient degree of correspondence.' Because statements with potential paranormal value refer most often to observable facts, the degree of correspondence between statement and fact was established by the researcher who was acquainted with the target person by applying the following criteria:

TABLE 2
Distribution of statements over the psychics

psychics	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
A:all	171	60	109	66	107	89	96	42	79	131	64	61
:ppv	19	3	4	2	6	8	17	4	3	20	3	6
:ppv %	11%	5%	4%	3%	6%	9%	18%	10%	4%	15%	5%	10%
B:all	124	69	110	81	90	124	88	39	55	128	73	65
:ppv	19	3	4	4	3	12	21	5	1	11	2	8
:ppv %	15%	4%	4%	5%	3%	10%	24%	13%	2%	9%	3%	12%

score 1: statement is untrue
score 2: statement is probably untrue or more untrue than true
score 3: statement is probably true or more true than untrue
score 4: statement is true

Statements of potential paranormal value which received a score of either 1 or 2 were called statements of negative paranormal value (untrue statements) and statements of potential paranormal value which received a score of 3 or 4 were called statements of positive paranormal value (true statements).

From table 3 it appears that the degree of correspondence between the content of the statements of potential paranormal value and the facts about the person is rather low. No difference was observed between the distributions of positive and negative paranormal statements for the two series A and B (Kolmogorov-Smirnov: Dmax= .06, Dcrit.01= .24). When one applies a cut-off criterion between the categories 'probably untrue' and 'probably true', both standard series together yielded 27 statements (14%) with positive paranormal value and 161 statements (86%) with negative paranormal value.

TABLE 3
Distribution of scores of 'degree of correspondence'

score	negative paranormal value 1	2	positive paranormal value 3	4
series A	54	30	10	1
series B	47	30	13	3
series A %	57%	32%	10%	1%
series B %	51%	32%	14%	3%
total A+B	101	60	23	4
total A+B %	54%	32%	12%	2%

The structural analysis of the sessions

(1) Topics discussed in the statements

Because the sitter asked for paranormal impressions without stating beforehand any specific problem area of the target person and because the sitter restricted himself as much as possible to a reactive type of behaviour, it was left to the psychic to decide which topics would be discussed and how much time would be devoted to each of them. The total number of statements in series A was 1075 and in series B 1046. Among other criteria for the itemization, a statement was considered finished when a new topic from the topic list was discussed (see for details Boerenkamp, 1984). As inclusive ('and'), exclusive ('or'), conditional ('if then') or causal ('because') verbalisations were excluded from this general rule, the number of topics discussed in the sessions is generally higher than the number of statements. For example: The person lives and works in Utrecht. This has been rated as one statement although it includes two topics (circumstances of work and circumstances of living). The total number of topics discussed by all psychics was 1423 for series A and 1433 for series B. The distributions of the topics discussed in the two series are presented in table 4.

TABLE 4
Distribution of topics in series A and B

Topics	A	B	A+B	A+B%	ppv
A Physical characteristics	(224	238	462	16%	73)
10 Sex	4	4	8	0%	0
11 Age	16	18	34	1%	5
12 Appearance and overt behaviour	66	65	131	5%	28
13 Bodily health	105	121	226	8%	27
14 Being alive or dead	33	30	63	2%	13
B Psychological characteristics	(645	579	1224	43%	45)
21 General personality traits	410	348	758	27%	26
22 Temporal psychological circumstances	192	206	398	14%	14
23 Religious and social orientation	43	25	68	2%	5
C Relations	(269	252	521	18%	26)
31 Relations with family members	111	112	223	8%	10
32 Relations with friends, colleagues	116	103	219	7%	11
33 Relation with sitter	42	37	79	3%	5
D Specific topics	(285	364	649	23%	93)
41 Civil status, number of children etc	42	36	78	3%	5
42 Profession, circumstances in work	93	157	250	9%	33
43 House, circumstances in living	39	55	94	3%	14
44 Leisure activity	57	60	117	4%	9
45 Specific name, property or event	54	56	110	4%	32

The distributions of topics in the two series is rather similar. The correlation between the frequency of the topics (subcategories 10-45) discussed in the two series is high (Spearman $r= .93$, $t=9.59$, $df=14$, $p< .0001$). This correlation is partly the result of the fact that the individual psychics appear to have consistent preferences for certain topics. If for each topic the distributions for series A and B are compared over psychics most correlations are significant. For physical characteristics: Spearman $r= .72$ ($t=3.31$, $df= 10$, $p< .01$), for psychological characteristics: $r= .81$ ($t= 4.39$, $df=10$, $p< .01$) and for specific topics: $r= .74$ ($t=3.46$, $df=10$, $p< .01$). The most variable theme appears to be the relations with other persons ($r= -.13$, $t= 0.40$, $df=10$, n.s.). The psychics appear especially consistent between sessions in the extent to which they discuss bodily health, being alive or dead, personality traits and specific names, properties, or events.

As might be expected, the distribution of the 237 topics (main categories A-D) discussed in the 188 statements with potential paranormal value (see column ppv in table 4) differs strongly from the distribution of the 2619 topics discussed in the other 1933 statements (chi-square= 107.05, $df=3$, $p< .0001$). Fewer descriptions of psychological characteristics and relations with other people are found in the set of statements with potential paranormal value whereas statements about physical characteristics and about specific topics are found relatively more often. The distributions of topics discussed in the statements with positive paranormal value and statements with negative paranormal value in terms of the main categories appear to be not significantly different (chi-square= 0.29, $df= 3$, n.s.).

(2) Person discussed in the topics

The psychics did not restrict themselves to statements concerning the target person. They also made statements about persons related to the target person (e.g. wife, father, child). Because of the itemization procedure, a specific statement might concern more than one person.
For example: The name of D or a family member of D is John.
Therefore, the number of topics rather than the number of statements concerning persons other than the target person were the basis for this evaluation.

In the 24 sessions the percentage of topics concerning persons related to the target person varied from 0% to 46% with a median of 6%. In series A a total of 153 of the topics (11%) concerned a person related to the target person. The total for series B was 204 (14%). This difference is significant (chi-square= 7.61, $df= 1$, $p< .01$). It further showed up that each psychic is rather consistent as regards the percentage of statements made about persons other than the target person. The correlation over psychics in the two series is positive and marginally significant (Spearman $r= 0.69$, $t= 2.97$, $df= 10$, $p<$

.02).

Finally, it appeared that the set of statements with potential paranormal value contained relatively more topics about a person related to the target person (26%) than the set of statements without potential paranormal value (11%) (chi-square= 51.46, df= 1, $p < .0001$). The distributions of statements of potential paranormal value with positive value and with negative value (i.e. the correct and incorrect statements of potential paranormal value) over the different types of persons discussed did not differ (chi-square= 0.26, df= 1, n.s.).

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

In the itemization procedure a change in the temporal character of a statement was not considered as the start of a new statement. Therefore, it is possible that one statement refers to circumstances from the past, present and future. In scoring each statement with respect to its temporal character the following rules were applied: (a) each statement in which a past and present circumstance was discussed was scored as a statement about the past and each statement in which a present and future circumstance was discussed was scored as a statement about the future; (b) in the (very exceptional) case that a past and future or a past, present, and future was discussed the statement was scored as a statement about the future. In table 5 the distribution of statements about past, present and future circumstances is presented.

The distributions of statements about past, present and future circumstances were significantly different for the two series (chi-square= 10.58, df= 2, $p < .01$). 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 (chi-square= 19.77, df= 2, $p < .0001$). Here again it appeared that psychics were rather consistent. The correlation between number of statements about the past of the target person in the sessions of series A and B was strongly positive (Spearman r= .84, t= 4.80, df= 10, $p < .0001$). The percentage of statements about the past in the 24 sessions varied from 0% to 39% with a median of 13%. However, there was no consistency in percentage of statements about the future (Spearman r= -.09,t= 0.28, df= 10, $p <$ n.s.). The number of statements

TABLE 5
Distribution of statements about past, present and future

	past		present		future	
	FH	SH	FH	SH	FH	SH
series A	82	99	423	392	33	46
series B	80	78	405	364	38	81
series A total	181		815		79	
series B total	158		769		119	
series A %	17%		76%		7%	
series B %	15%		74%		11%	
total FH-SH	162	177	828	756	71	127
total A+B	339		1584		198	
total A+B %	16%		75%		9%	

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

about the future depended partly on the specific circumstances of the individual target person. The percentage of statements about the future in the 24 sessions varied from 0% to 30% with a median of 10%.

Finally, 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 future (15%) compared to the set of statements with no potential paranormal value (chi-square= 27.29, df= 2, $p < .0001$). The distributions of the statements of potential paranormal value with positive and negative value related to the past, present and future did not differ (chi-square= 3.67, df=2, n.s.).

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

Usually people turn to psychics because of problems they have. Therefore, it can be expected that psychics are inclined to discuss more often such problem areas even when they are not indicated beforehand. In scoring each statement with respect to its positive or negative emotional character the following rules were applied: (a) a statement in which only a fact was stated was scored as neutral (for example: his name is John); (b) a statement in which a favourable as well as an unfavourable state of affairs was discussed was scored as neutral (for example: he has difficulties now but it will be less in a couple of years); (c) a statement about a personality trait was scored according to its social desirability (for example: 'he is stubborn' was scored as unfavourable and 'he is intelligent' was scored as favourable). Table 6 presents the distribution of statements concerning a favourable, neutral, or unfavourable state of affairs.

TABLE 6
Distribution of statements concerning a favourable, neutral, or unfavourable state of affairs

	favourable		neutral		unfavourable	
	FH	SH	FH	SH	FH	SH
series A	83	97	178	186	277	254
series B	115	124	183	206	225	193
series A	180		364		531	
series B	239		389		418	
series A %	17%		34%		49%	
series B %	23%		37%		40%	
total FH-SH	198	221	361	392	502	447
total A+B	419		753		949	
total A+B %	20%		35%		45%	

The distribution of the statements in the two series appeared to differ to a significant degree (chi-square= 22.20, df= 2, $p < .0001$). Series B yielded relatively more statements of a favourable nature. When the sessions are split up in first and second halves, it appeared that statements about an unfavourable state of affairs slightly preponderated in the first half of the session (chi-square= 5.73, df= 2, $p < .06$). It also appeared that psychics were rather consistent in the percentage of statements they made about an unfavourable state of affairs (Spearman r= .63, t= 2.56, df=10, $p < .05$). However, this did not hold for statements about a favourable state of affairs (Spearman r= .28, t= 0.93, df=10, n.s.). This finding suggests that the number of statements about a favourable state of affairs partly depends on the specific circumstances of the target person. The percentage of statements about a favourable state of affairs in the 24 sessions varied from 7% to 61% with a median of 18%. The percentage of statements about an unfavourable state of affairs varied from 3% to 87% with a median of 39%.

Finally, it appeared that the set of statements with potential paranormal value 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 distributions of statements with positive and statements with negative paranormal value indicating a favourable, neutral or unfavourable state of affairs did not differ either (chi-square= 4.81, df= 2, n.s.).

(5) Number of statements in the form of advice

Occasionally a psychic gives advice to the target person; for instance, when he states about a friend of the target person that it wouldn't be a loss if that person disappeared from the target person's life because the psychic feels that this friend is not reliable. In series A, 53 statements (5%) involved advice. In series B, it was 67 statements (6%). The difference is not significant (chi-square= 1.89, df= 1, n.s.). The psychics appeared to offer a consistent percentage of statements involving advice. The correlation between the two series is positive (Spearman r= .73, t= 3.36,df= 10, $p < .01$). The percentage of statements in the form of advice varied from 0% to 17% with a median of 3% in the 24 sessions. It has to be noted that the differences between the psychics in this respect were very strong. Two of the psychics were responsible for half of all statements offering advice, while three did not give any advice at all.

Finally, it appeared that compared to the other statements the set of statements with potential paranormal value contained fewer statements involving advice (3% versus 6%). The difference is marginally significant (chi-square= 4.05, df= 1, $p < .05$). Because the total number of statements with potential paranormal value involving advice is not more than 4 it is impossible to compare statistically the set of statements with positive paranormal value and the set of statements with negative paranormal value with respect to this characteristic.

(6) Number of times a silence precedes a statement

Now and then during a session a psychic remains silent. It might be that in these periods the psychic experiences paranormal impressions more intensively. If so this might affect the statements which follow such a period of silence, resulting in a higher percentage of statements with potential paranormal value and with positive paranormal value. A silence was defined as a pause of 3 seconds or more and was indicated in the transcripts. In series A, 162 statements (15%) were preceded by a silence and in series B, 175 statements (17%). The difference between the two series is not significant (chi-square= 0.97, df= 1, n.s.). The psychics appeared to have a very consistent style in making a pause in the session. The psychic who made relatively more pauses in the session of series A tended to to this again in the session of series B (Spearman r= .90, t= 6.56, df=10, $p < .0001$). The percentage of statements preceded by a silence varied from 1% to 59% in the 24 sessions, with a median of 12%.

Finally, it appeared that the number of statements with potential paranormal value did not differ from the set of statements without potential paranormal value in this respect (chi-square= 1.39, df= 1, n.s.). The set of statements with positive value and the set of statements with negative value also did not differ with statements preceded or not preceded by a period of silence (chi-square= 1.99, df=1 , n.s.).

(7) Number of positive and rhetorical statements

In both series about 30% of the statements of the psychic were rhetorical statements. In rhetorical statements the psychic asks for immediate feedback in such a way that the sitter is invited to react at least with either "yes", "no", or "don't know". Such statements are either statements made in question mode or statements followed by the remark "Can you 'place' that?" A significant difference appeared between the first and second half of the sessions as regards the number of rhetorical statements, such that more rhetorical statements were found in the first half of the sessions (chi-square= 14.20, df= 1, p< .001). In table 7 the distributions of rhetorical and positive (non-rhetorical) statements for series A and B are presented.

TABLE 7
Distribution of positive and rhetorical statements

	positive			rhetorical		
	FH		SH	FH		SH
series A	348		389	190		148
series B	352		391	171		132
series A		737			338	
series B		743			303	
series A %		69%			31%	
series B %		71%			29%	
total FH-SH	700		780	361		280
total A+B		1480			641	
total A+B %		70%			30%	

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

This result implies that psychics invite feedback more often in the first part of the session. A comparison based on the first 10%, middle 10%, and last 10% of the statements in each session confirmed this conclusion (chi-square= 47.5, df= 2, p< .0001). Here again it was found that the individual psychics were rather consistent in the percentage of rhetorical statements. The correlation between the sessions of series A and B over psychics is positive (Spearman r= .67, t= 2.86, df= 10, p<.02). The percentage of rhetorical statements in the 24 sessions varied from 5% to 54% with a median of 27%.

Finally, it appeared that the set of statements with potential paranormal value contained more rhetorical statements (65%) than the set of statements with no potential paranormal value (27%). This difference is strongly significant (chi-square= 112.24, df= 1, $p<.0001$). As regards positive and rhetorical statements the set of statements with positive paranormal value did not differ from the set of statements with negative paranormal value (chi-square= 0.01, df=1, n.s.).

(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 8. If the psychic made two statements before the sitter reacted it is indicated by P2S etc..

TABLE 8
Distribution of the statements preceding an informative action

	P1S	P2S	P3S	P4S	P5S	P6S	P7S	P8S	P9S	P10S	P11S
series A	493	136	43	13	10	5	2	2	1	1	0
series B	471	137	41	18	8	6	2	2	0	0	0
total	964	273	84	31	18	11	4	4	1	1	0

The distributions observed in the two series are nearly equal (chi-square= 1.76, df= 6, n.s.). (Categories P7S to P11S combined). The sitter reacted immediately after about 45% of the statements. It appeared that the psychics as well as the sitters were rather consistent in this respect. The psychics who elicited most P1S interactions (in terms of the percentage of all types (P1S to P11S)) in the sessions of series A tended to do this again in the sessions of series B (Spearman r= .73, t= 3.37, df=10, $p<.01$). The percentage of

PlS type of interaction in the 24 sessions varied from 1% to 83% with a median of 43%.

Finally, it appeared that the set of statements with potential paranormal value contained on the average the same number of statements which were part of PlS interactions as the set of statements with no potential paranormal value (chi-square= 0.02, df= 1, n.s.). The set of statements with positive paranormal value and the set of statements with negative paranormal value did not differ in this respect either (chi-square= 0.21, 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. In series A the sitter denied the correctness of 92 statements. The comparable value for series B is 72 statements. This difference between the two series is not significant. From a content analysis it appears that the psychics mainly used four types of responses to a denial:

(1) Accepting the denial by giving another impression (AD).

For example: P: Her father is still alive (?)
S: No (--), he died some years ago
P: No
Silence: 5 sec
P: Is she ill (?)

(2) Giving a new interpretation to the denied impression (NI).

For example: P: She does not live together with her friend (?)
S: No (--), she does
P: They separated once in the past (?)
S: No (--), they didn't
P: I think they will divorce in the future as she is a jealous type of person

(3) Suggesting the target person knows better (T>S).

For example: P: Does he likes singing because I hear songs (?)
S: No I don't think so (-?) as I never noticed it
P: You might better ask as he does like it

(4) Suggesting the content of the informative action is equal to the content of the statement (I=S).

For example: P: Is the person over 50 years already (?)
S: No (--) he is 45

P: I would say around 50 and then he must take
care of being promoted.

In table 9 the different responses of the psychic are presented.

TABLE 9
Distribution of types of responding to a denial

	AD	NI	T>S	I=S
series A	12	57	15	8
series B	10	44	10	8
total	22	101	25	16
total %	14%	61%	15%	10%

The difference in the distribution among the different types of reactions in the two series is not significant (chi-square= 0.42, df= 3, n.s.). It appeared that in a majority of these types of interactions the psychics (86%) use some defensive reaction. The low numbers do not permit an analysis in terms of the consistency among certain types of responses for each psychic.

DISCUSSION

From the results of the informational analyses of these series it appears that the effects of the assumed specific paranormal ability of psychics on the statements is rather limited if not absent. Before a paranormal character is attributed to a set of statements, the criteria of 'no logical explanation' and 'sufficient degree of correspondence' have to be met. It was found that approximately one out of ten statements can be considered to meet the criterion of 'no logical explanation'. Furthermore, on average only one out of ten of these statements meet the criterion of 'sufficient degree of correspondence'. Probably none of the 1% of statements which meet both criteria can be considered to have a probability of correspondence lower than 1 in 100. Therefore it seems that the observed 1% 'unexplainable' correspondences between the statements and the facts concerning the target person can be satisfactorily explained by assuming chance coincidence.

As stated in the design (see Boerenkamp, 1985, this issue) a more definite conclusion about the assumed paranormal ability of psychics has to be based on all sessions of the investigation. However, assuming that our tentative conclusion is correct, it is not surprising that in most analyses discussed above no differences were found between the distributions of statements with potential paranormal value and statements without potential paranormal value. In addition, in those analyses where differences were observed they can be considered 'logical' consequences of the estimation procedure. For instance, the preponderance in statements of potential paranormal value of descriptions of specific names, properties or events, of circumstances in work and living and of bodily health is not surprising because these topics are often more specific compared to descriptions of personality traits and temporal psychological circumstances. The preponderance of statements about persons related to the target person is an indirect result of it. If psychics talk about persons related to the target person they are less likely to concentrate on the psychological characteristics of these persons and their relations with other people. In the set of statements with potential paranormal value the preponderance of statements related to the past of the target person has probably two causes. One is the relatively low number of unspecific statements about personality traits in this class. Such statements, which were all classified as 'present', are generally rated of no potential paranormal value. The other reason is the relatively low number of statements about the future in this class, which are often based on the preceding interaction concerning the present state of affairs and therefore are more often judged as 'inferred'. Consequently such statements also have a higher probability of being classified as having no potential paranormal value. The relatively low number of statements involving advice in the set of statement with potential paranormal value can be explained in a similar way. Such statements usually follow some interactions and therefore are more often classified as 'inferred'.

The only difference between statements with potential paranormal value and the others which seems not to be a 'logical' consequence of the judging procedure concerns the strong preponderance of rhetorical statements in the set of statements with potential paranormal value.

The explanation for this finding might be that the psychic invites immediate feedback when he verbalizes a really specific or spontaneous impression in order to reduce the concommitant uncertainty as much as possible. As observed before, the psychic is expected to make statements with a low probability of correspondence. This result suggests that the psychic is aware when he takes a risk.

The distribution of the different characteristics of the statements over the individual psychics strongly indicate that the psychics appear to be very consistent, if not stereotyped, in their behaviour. Each of them has a consistent preference for certain topics, for assigning a certain percentage of statements to persons related to the target person, for assigning a particular percentage of statements to the past, for assigning a particular percentage of statements to unfavourable states of affairs, for giving a particular percentage of statements in the form of advice, for using a particular number of silences in the session, for making a particular percentage of rhetorical statements, and for employing a consistent style in eliciting immediate feedback in general. In short, they show consistent behaviour in most characteristics studied. Moreover, they have a very consistent preference for the duration of a session. In fact, the particular target person and his or her specific circumstances of life only influence significantly the number of statements about relations with other persons, the number of statements about the future and to a lesser degree the number of statements about favourable states of affairs.

Considered as a group, the psychics also appear rather consistent in their behaviour. The number of statements in both series is virtually equal. The same applies to the distribution of all characteristics except one. The only strong inconsistency between the two series is the relative difference in number of statements about favourable and unfavourable states of affairs. To a certain degree this number is influenced by the different ways the psychics perceive the target persons and their circumstances of life, i.e. psychic 1 has a positive impression of target person 1 and psychic 2 has a negative impression of the same target person.

Some slight differences were found between the first and second halves of the sessions. The statements with potential paranormal value were found in the first halves rather than in the second halves in both series but the difference appeared not to be significant for all sessions. Furthermore, a slight preponderance of statements about an unfavourable state of affairs was found in the first halves. Possibly this indicates that psychics expect sitters to indicate a problem of the target person at the beginning of a session. A strong preponderance of statements about the future was found in the second halves of the sessions. In the same vein, this might indicate that in the daily routine, towards the end of the session sitters expect some advice concerning their future circumstances.

The interactional analysis might indicate first of all that the psychics felt themselves rather 'safe' with the sitters. They probably experienced a permissive atmosphere for making 'mistakes', as they simply gave a new interpretation to the denied impression in the majority of the cases. In this respect the results of the interactional analysis might deviate from what would have been found in normal daily routine. It will be of interest to compare the standard series with series 10, in which the target person was present. It is conceivable that other categories of reaction to denial will have to be added to deal with such an 'interactionally' richer situation. In the standard series, only in 10% of the cases on the average in which the sitter denied the correctness of the content of a statement made by the psychic did the latter risk a real disagreement by suggesting that the denial was unjustified.

ABSTRACT

The aim of the study presented in this paper was to provide a description of the content of sessions when psychics are active under conditions which resemble to a large degree the daily circumstances of sessions with clients. Two series of sessions were held with a group of 12 psychics in which they were consulted about different target persons. The psychics received feedback as normally is the case in their daily routine.

On the average about one of ten statements appeared to have a medium-high or high potential paranormal value. Of these statements about 10% were found to be correct. Hence only about 1% of all statements can be considered as possibly based on paranormal impressions. In view of the judging procedures applied, this finding is not supportive of a paranormal interpretation.

From a psychological point of view one of the most interesting

results of the analyses of the standard series is that psychics appear to be very consistent in their behaviour. Each psychic has a rather consistent preference for sessions of a certain duration. There is a strong correlation between the length of sessions expressed in number of statements of the different psychics between the two series. The psychics as a group produce nearly the same number of statements in both series. Psychics appear to be consistent in most other characteristics studied. It also appears that the target person and his particular circumstances in life hardly affects the structure of the verbal behaviour of the psychics. These findings indicate that the verbal behaviour of psychics in sessions in which their statements are supposedly based on paranormal impressions can be considered rather stereotyped.

ACKNOWLEDGMENT

The author gratefully acknowledges his indebtedness to Dr S.A. Schouten for his critical appraisal of the design and his willingness to participate as the co-researcher in the sessions.

REFERENCES

Bender, H. 'Praekognition im qualitativen Experiment', Zeitschrift fur Parapsych. und Grenzgeb. der Psych., Bd I, 1957, 5-35.

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 aquired under feedback conditions', E.J.P., 5, 2, 1984, 101-124.

Hettinger, J. 'The Ultraceptive Faculty', London, Rider, 1940.

Hettinger, J. 'Exploring the Ultra-Perceptive Faculty', London, Rider, 1941.

Nilsson, I. 'The paradigm of the Rhinean School, Part I', E.J.P., 1, 1, 1975, 45-59.

Palmer, J. 'Extrasensory perception: Research Findings' in 'Advances in Parapsychological Research, vol 2, Extrasensory Perception', Ed. S. Krippner, Plenum press, New York, 1978, 59-243.

Palmer, J. 'ESP Research Findings: 1976-1978' in 'Advances in Parapsychological Research, vol 3', Ed. S. Krippner, Plenum Press, New York, 1982.

Pratt, J.G. 'Towards a Method of Evaluating Mediumistic Material', Bull. Boston SPR, 23, 1936.

Pratt, J.G., Birge, W.R. 'Appraising Verbal Test Material in Parapsychology', J.o.P., 12, 1948, 236-256.

Pratt, J.G. 'On the Evaluation of Verbal Material in Parapsychology', Parapsychology Monographs no 10, Parapsychology Foundations Inc, New York, 1969.

Rhine, J.B. 'Extra-sensory Perception', Boston, Bruce Humphries, 1934.

Roll, W.G. 'Designs for tests with free response material', J. A.S.P.R., 56, 1962.

Schmeidler, G. 'Analysis and evaluation of proxy sessions with Mrs. Caroline Chapman', J.o.P., 22, 1958, 137-155.

Schouten, S.A. 'De standaardmethode als instrument voor Parapsychologisch Onderzoek', Utrecht, 1973, dissertation.

Scott, C. 'Experimental Object-Reading: A Critical Review of the work of Dr J. Hettinger', Proceedings of the SPR, vol. XLIX, 1949, 16-50.

Scott, C. 'On the Evaluation of Verbal Material in Parapsychology: A Discussion of Dr Pratt's Monograph', J.S.P.R., 46, 1972, 79-89.

Tenhaeff, W.H.C. 'Experimentele proscopie', Tijdschrift voor Parapsychologie, Bd XXI, 1953, 49-109.

Tenhaeff, W.H.C. 'Proceedings of the Parapsychological Institute of the State University of Utrecht', number 1 (1960), number 2 (1962), number 3 (1965).

Tenhaeff, W.H.C. 'Ontmoetingen met paragnosten', Bijleveld, Utrecht, about 1980 (undated).

Timm, U. 'Die statistische Analyse Qualitativer Paragnostischer Experimente', Zeitschrift fur Parapsych. und Grenzgebiete der Psych., Bd XXI, 1965, 78-122.

EUR.J.PARAPSYCHOL., V.5, 1985

A RETRO AND DIRECT PK TEST FOR BABIES WITH THE MANIPULATION OF FEEDBACK: A FIRST TRIAL OF INDEPENDENT REPLICATION USING SOFTWARE EXCHANGE

DICK J. BIERMAN

INTRODUCTION

According to the Observational Theories, coupling between the psi source and the random system occurs when the psi source consciously observes the behaviour of the random system. On the one hand, this basic assumption constrains the region of action of the otherwise omnipotent psi mechanism. On the other hand, combined with the assumed fundamental space-time independent character of this coupling, future observers of the experimental results might be psi sources in our experiments, which would explain the rather unreliable experimental results.

However in the history of psi research there have been outstanding and rather reliable subjects (and/or experimenters) who obviously could withstand the supposed 'noise' of a large number of future psi sources. This indicates that if we could find a very strong psi source, like those of the old days, it would be possible to do some fundamental theory testing that might be fairly reliable.

It has been found by Spinelli (1976) that very young children of about 3 years of age show extremely high scoring rates in psi tests, significantly higher than any other age group. In the present experiment babies of about 10 months were tested in the hope that scoring rates associated with these subjects would be even higher.

With these subjects, two fundamental hypotheses of the Observational Theories were tested: namely, the effect of FEEDBACK (Observation) and the effect of PRERECORDING of trials (Time independency). Furthermore, the design was such that the observational history of the data, including the ANALYZING stage, was completely formalized and manipulated.

According to most variants of the Observational Theories, the 'time dependency' implies that the experiment is not really finished after the subjects 'have left the stage'. The results are assumed to be dependent also on the 'psi input' of further observers of the data (e.g. the analyzer of the data). In the hierarchical version of the Observational Theories (Houtkooper, 1980, 1983) the order of the observation is relevant because each observer is considered not only to be a potential source of psi but also as a 'psi sink', for the psi of subsequent observers. Research lines exploring the effect of preobservation of data (Weiner and Bierman, 1981; Bierman and Weiner, 1981; Schmidt, 1983) are based upon similar notions. For instance another variant of the Observational Theories holds that only the first observer can be the psi source, albeit that further psi input is allowed, either due to forgetting of the first observer or to observation of data related to but not quite dependent of the first observed data (e.g., the first subject observes trial by trial scores while the analyzer looks at runscores split for sex).
In the present experiment part of the data is not fed back to the subject. According to the hierarchical version as well as to the first observer only version of the Observational Theories, this would increase the efficiency with which potential analyzer psi would bias the data.

The replicability issue has played a consistent role in discussions about the reality of psi phenomena. Although, in principle, replicability is no conditio sine qua non for phenomena to be accepted as real, it is of course a conditio sine qua non to advance parapsychology as an empirical science.

Non-replicability traditionally has been explained by uncontrolled variance due to unknown variables. For instance, the rationale behind animal research was that probably fewer variables might be relevant and hence replicable results could be expected. Some authors claim that within-laboratory replicability is already obtained in some parapsychological research centers (Sargent, 1982), and the between-laboratory non-replicability is explained by the fact that written reports are not adequate to communicate all the relevant conditions. This latter has for example been shown with regard to the failure to reproduce the early results with Lasers (Collins, 1978).

Near the end of the 1970's Honorton (1978) proposed to establish a network of research centers equipped with small computers, based on the argument that such an approach might bring about between-laboratory replicability because the experimental procedure would be completely formalized and could be communicated by exchange of software. Although this idea was formulated more than 5 years ago, and although in the meanwhile an 'APPLE PSI USER GROUP' has even been established, there have been no substantial efforts published documenting the feasibility of this approach. The idea to formalize the experimental procedure through the use of micro-computers has become even more attractive since the breakthrough of personal computers can put these on the desk of skeptics.

Apart from the fact that the present experiment provided for such an independent replication because it is completely controlled by a computer program, there are two additional arguments for running it.

In the first place, the subject pool is well defined and (using similar arguments as for the animal subject pool) might contain less variance. In the second place, there are theoretical arguments that would predict better replicability in the present experiment. The Observational Theories identify the unknown variables which 'cause' non-replicability as being the future observers of the data, such as experimenters and analyzers. In the present experiment, at least one part of the future, namely the data analysis, is completely formalized and controlled.

Therefore, it was decided that after the first series (the Amsterdam series), a second series (The Bristol series) would be conducted as a replication and as an attempt to explore the procedures that are needed to develop a between-laboratory replication procedure following the lines sketched originally by Honorton.

THE AMSTERDAM SERIES

Experimental procedure

The experiment consisted of 24 runs of attempted PK on a high speed RNG connected to a microcomputer. Two subjects, aged 10 months, both related to the Experimenter, participated. The 4096 trials of each run were composed of four-trial blocks as follows:

- The first trial was instantaneously generated and resulted in feedback if it corresponded to the target.
- The second trial was instantaneously generated as well but did not result in feedback.
- The third trial was prerecorded and resulted in feedback if a hit occurred.
- The fourth trial was also prerecorded but did not yield feedback.

The following blocks of four trials were composed in the same way, thus resulting in 4 groups of 1024 trials each per run. The RNG produced numbers in the interval 0 to 255. Before each run was initiated, 2048 samples of the RNG were prerecorded in the memory of the computer. Then a random target was generated from the same interval (0,255).

In the meantime the subject was prepared and seated in the baby chair about .5 meters in front of a video display. Fifteen seconds before the actual start of the run, a meaningless block pattern appeared on the screen, while a countdown beep was used to focus the attention of the subject on the screen and away from the experimenter who was leaving the room. After countdown the run started, with hits occurring whenever a trial corresponded to the target ($p=1/256$). For the feedback condition (odd trials) a hit resulted in a change of the displayed block-pattern into a laughing face and a melody was played for 3 seconds. For the non-feedback condition (even trials) the hits were only counted. Each run lasted about 2 minutes, during which time 4096 trials, including the 2048 prerecorded ones, were compared with the target. The experiment took place during the months of April through July, 1982, with generally 3 runs per experimental day and one experimental day per week.

Randomness tests

As a daily routine, the RNG used in this study was tested before, during and after the experiment for first till 8th order non-randomness. Test runs consisted of 100000 samples sampled at the highest possible sampling rate using another program. There were no consistent patterns of non-randomness found during these tests.

Analyzing procedure

After completion of all 24 runs, the experimental results, which were recorded on disk, were split into 2 parts using an ABBA scheme based upon run number.

The 2 parts were informally analyzed independently by two analyzers (A=DJB and B=JMH, alphabetical assignment) for the effect of FEEDBACK and PRERECORDING using 2 sample t-tests. This specific pair of analyzers was selected because they had obtained analyzer effects in previous work (Bierman & Houtkooper, 1980). They therefore can not be considered as random variables. Their specific goal was to obtain conflicting results (e.g. Analyzer A finds a feedback effect with hitting in the feedback condition and analyzer B finds the opposite).

The data for both subjects were pooled since the subject is not a variable of interest in this study.

As the unit of analysis, the sessionscore was used. Since the p-value for a hit is extremely small, and the number of trials per session was not extremely large, the distribution of these scores might deviate from the normal distribution.

HYPOTHESES

Hypothesis 1

There will be a differential effect between feedback and non-feedback trials.

Hypothesis 2

There will be an Analyzer effect. It will be most pronounced in the non-feedback trials.

Hypothesis 3

There will be no difference in scoring rate between prerecorded and instantaneously generated trials.

Although originally an ANOVA was planned, the possible non-normality of the score distribution forced us to test these hypotheses by means of a 2 by 2 Chi-square test. It should be remarked that this test does not yield pure estimates since it does not take into account interaction effects.

Hypothesis 3 will be tested with a p of 0.10 since it would be easy to miss a real difference due to the difference being small and a rather small sample size.

A confirmation of hypothesis 1 would support the basic assumption of all Observational Theories, namely that feedback is a necessary condition for psi to operate.

A confirmation of hypothesis 2, would support the Observational Theories that assume an order of observation dependent on the 'addition rule'.

A confirmation of hypothesis 3, would support the assumption of 'time independency'.

RESULTS

After the completion of the experiment both analyzers noted extreme below chance scoring on the prerecorded sets. It turned out that a last-minute change in the program that was necessary to clear the graphics screen also was responsible for the clearance of a part of the prerecorded set. Therefore the number of trials per run in the prerecorded conditions was not the planned 1024 but 473 (feedback) and 472 (non-feedback). To maintain the possibility of comparing the raw data in each condition (APPENDIX I), these are therefore expressed as continuity corrected z-scores in TABLE 1.

TABLE 1
z-scores for all conditions (Amsterdam series)

	FEEDBACK to babies	NO-FEEDBACK to babies	INSTANT. trials	PREREC. trials	total
Analyzer DJB	-0.66	-2.09*	-1.41	-1.50	-1.98*
Analyzer JMH	-0.07	-0.15	+0.69	-1.40	-0.18
Total:	-0.55	-1.61	-0.43	-2.08*	-1.55

The two interesting figures (designated *) are those for the overall missing in the data analyzed by analyzer DJB (z=-1.98) and the overall missing in the prerecorded condition (z=-2.08). The formal test of the hypotheses by the Chi-square (APPENDIX II) showed however that none of these could be confirmed.

DISCUSSION OF THE AMSTERDAM SERIES

If the below chance scoring in some of the conditions is due to psi, then we are faced with the problem of the interpretation. The suggestive overall missing in the data analyzed by analyzer DJB seems to point to an analyzer effect. The data of analyzer JMH however also showed (non-significant) missing and hence the hypothesis 2 concerning the Analyzer difference effect could not be confirmed.
Note that the effect of analyzer DJB is strongest in the data that were not observed by the babies.

It should be remarked that the difference between INSTANTANEOUSLY and PRERECORDED trials, far from being significant, could be due to confounding with a within run decline effect. The prerecorded trials were only presented during the first half of the run since the latter part of the prerecorded data had been erased inadvertently.

THE BRISTOL INDEPENDENT REPLICATION SERIES.

Experimental procedure

The experiment was planned for 24 runs. To have the same total number of trials as in the Amsterdam series, the number of trials per run was doubled, with a corresponding increase in trial rate. However the BRISTOL experimenter had understood that 48 runs were necessary instead of the planned 24. The experiment was stopped after 33 runs on the request of the present author, who had no information on the accumulated results, as soon as this communication problem was discovered.
The same computer program that was used in the Amsterdam series controlled the Bristol series, apart from one internal modification that was made to enhance possible theoretical relevance. The prerecorded trials in the Bristol series were presented 4 times within the run. According to certain versions of the Observational Theories this should increase the scoring rates in the prerecorded condition (Schmidt, 1976). From the subject's point of view both series were completely indiscriminable.

Randomness test

Randomness tests using a special program were performed before, during and after the series. No consistent patterns of non-randomness were detected.

Analyzing procedure

After the experiment, the Bristol experimenter, who is known to be very sceptical about the reality of psi phenomena, returned the printouts and the diskette that contained the program and a special logfile to Amsterdam. The logfile was not readable nor modifiable by a non-sophisticated computer user. It contained some hidden code so that even in the case of breaking into the file this would become obvious on inspection. Whenever a run was initiated, the date, the time and the name of the subject were written on the logfile. Upon completion of the run the results were written on the file. The returned printed output could be checked against the logfile data to look for abnormally terminated runs. However, before any observation of the data was done, the printed output was divided among the same analyzers as in the Amsterdam series, again using an ABBA splitting scheme based on the order of the printouts in the returned package. Upon termination of this first stage of the analyzing procedure, which was identical to the one in the Amsterdam series, analyzer DJB checked the logfile against the printouts and found two discrepancies. In the first place, two printouts were missing. The relevant data, which were available in the logfile, were thus added to the data pool for further analysis. In the second place the logfile indicated that one run of the subject was terminated before the run had been finished. It turned out that this had occurred due to malfunctioning of the computer with loss of data.

HYPOTHESES

The hypotheses 1 and 2 were identical with the Amsterdam-series hypotheses.

No formal predictions were made with regard to the effect of the multiple presentation of prerecorded trials.

RESULTS

The results in terms of z-scores corrected for continuity are given in TABLE 2.
In contrast to the Amsterdam series, the overall results are in the hitting direction. The hitting is concentrated in the FEEDBACK trials. Compared to the marginal results of the Amsterdam series the result of the Bristol series is quite straightforward. The formal analysis of the FEEDBACK effect (APPENDIX II) yields a significant Chi-square of 5.59 ($p<.02$). The exploration of the effect of multiple observation of the prerecorded data shows null-results.

TABLE 2
z-scores for all conditions (Bristol series)

	FEEDBACK to babies	NO-FEEDBACK to babies	INSTANT. trials	PREREC. trials	total
Analyzer DJB	+1.65	0.00	+1.00	+0.30	+1.17
Analyzer JMH	+1.94*	-1.22	0.00	+1.07	+0.48
Total:	+2.56*	-0.85	+0.72	+1.00	+1.21

DISCUSSION OF BRISTOL AND AMSTERDAM SERIES

The results of the two series differ in two respects:

1: In the first place there was psi missing in the Amsterdam series while there was hitting in the Bristol series. Although similar reversals from pilot to confirmation series are not uncommon in the literature, there have been no satisfactory 'explanations' put forward to account for this apparent negative reliability (Bierman, 1980). In the present case such a (pseudo) 'explanation' would be that the subject(s) in the first study disliked the reinforcement that was contingent on a hit while the subject in the second series liked the very same reinforcement. The use of two-tailed tests might justify conclusions in evidential research but tends to obscure a fundamental aspect of psi in process-oriented research. I have speculated elsewhere that such reversals might be conceptualized as a 'relaxation' effect. In virtually every physical system that is in equilibrium some disturbance of the system results in a tendency of the system to return to its equilibrium state sometimes showing some 'overshoot'. The 'agent' that is responsible for the relaxation in psi disturbed systems might be the collective scientific community. Such a speculation transfers the phenomenon of psi missing from the traditional domain of psychological 'explanations' to the domain of observational theoretical 'explanations'.

2: In the second place, the Amsterdam series did not confirm the main hypothesis concerning the manipulation of Feedback while the Bristol series did. The Feedback related hypothesis is of course based upon one of the essential assumptions of the Observational Theories. Could we conclude, therefore, that the Amsterdam series did not support the theory? In the discussion of the Amsterdam series, it was indicated that such a conclusion is not justified but that the rather marginal results of the Amsterdam series, assuming that there was psi in the data, might tentatively be ascribed to psi of analyzer DJB. However, the results of the Bristol series are consistent over both analyzers. Although this does not necessarily imply that there was no analyzer psi input, it supports an interpretation in terms of the subject being the major psi source. A possible 'explanation' for this screening of analyzer psi could be that, although the Bristol experimenter did not analyze the results, she occasionally observed the printouts and extracted some meaningful information in terms of analysis. (She literally said: "I think there is something in the data"). This unintended observation on the part of the experimenter might have constrained the effects of further future observers, i.e., the analyzers.

DISCUSSION OF THE INDEPENDENT REPLICATION PROCEDURE

Since experimental parapsychology is a 'statistical' science by nature, the building up of knowledge is a slow process that requires data of many similar experiments. For instance the Observational Theories can not be falsified and even less 'proved' by a single experiment. Therefore the procedure of 'replication by software exchange' might have a significant impact, enabling meta-analysis of well defined datasets across laboratories.

The failure in communication concerning the total number of sessions suggests that in further experiments along these lines the software should keep track of the number of sessions. This might be done through the protected Logfile (see B: Analyzing procedure), which proved to be an essential 'tool' in this approach. On the other hand, there is the drawback that by eliminating all experimenter responsibilities the motivation and involvement of the independent replicator might become too low to even finish the series. At present, the necessary cooperation (which is needed from a group of independent replicators for each project) can only be realized on a reciprocal basis.

ABSTRACT

Two babies of about 10 months of age were subjects in the first series, while a single subject of the same age was used in a completely independent replication series. During the experimental session the subject was seated in front of a visual display that was controlled by a computer with a Random Number Generator. Before each session a target number was generated and a series of numbers from the RNG was prerecorded in the memory of the computer (PRERECORDED condition). During the session, instantaneously generated numbers as well as prerecorded numbers were compared with the target number. When a hit ($p=1/256$) occurred the nonsense block pattern display was replaced by a display of a laughing face and a melody played for a few seconds. However, for the even 'trials' only the hits were counted and the display was not changed (NO-FEEDBACK condition).
The results of the sessions were split in two parts and analyzed by two different analyzers (ANALYZER condition). For the first (Amsterdam) series overall scoring was in the missing direction ($z=-1.55$), mostly due to the data analyzed by analyzer A ($z=-1.98$, $p<.05$). The main hypothesis, predicting a FEEDBACK effect, was not confirmed.
The computer program controlling the experimental sessions was mailed on diskette to a 'sceptic' who performed the second series completely independently. The data of this second (Bristol) series were returned to be analyzed in Amsterdam. They confirmed the main hypothesis showing a significant feedback effect (Chi-square=5.59, df=1, $p<.02$) with significant hitting in the FEEDBACK condition ($z=2.56$, $p<.015$).

ACKNOWLEDGMENTS

I want to thank Dr. Sue Blackmore for her contribution as experimenter in the second series. I want to thank Dr. Joop Houtkooper for the afternoons of competitive analyzing.

REFERENCES

Bierman, D.J. 'Negative reliability: The ignored rule', in W.G. Roll and J. Beloff (eds.), Research in Parapsychology 1980, Metuchen, N.J., Scarecrow Press, 1981, 14-15.

Bierman, D.J. and Weiner, D.H. 'Multiple preobservation of data as a method for screening off future-observer effects', in W.G. Roll, R.L. Morris and R.A. White (eds.), Research in Parapsychology, 1981, Metuchen, N.J., Scarecrow Press, 1982, 132-134.

Collins, A. 'A sociological study of the scientific method', Proc. of symposium: Replication and experimenter effects, A.A.A.S., Feb. 1978, Washington.

Honorton, C., Personal Communication, 1978.

Houtkooper, J.M., Andrews, R., Ganzevles, P.G.J. and Sijde, P.C. v.d. 'A hierarchical model for the Observational Theories: A study of subject identity and feedback in repeated retroactive psychokinesis', E.J.P., 1980, 3, 221-246.

Houtkooper, J.M. 'Observational Theory: A research programme for paranormal phenomena', Thesis, 1983, Publ: Swets & Zeitlinger, Lisse.

Sargent, C.L. 'Replicomania and the pursuit of meaning in experimental parapsychology', In B. Shapin and L. Coly (eds.), Parapsychology's second century, New York, N.Y., Parapsychology Foundation, 1982, 99-118.

Schmidt, H. 'PK effect on pre-recorded targets', J. A.S.P.R., 70, 1976, 267-291.

Schmidt,H. 'Superposition of PK efforts by man and dog', in R.A. White and R.S. Broughton (eds.), Research in Parapsychology 1983, Metuchen, N.J., Scarecrow Press, 1984, 96-98.

Spinelli, E. 'The effects of chronological age on GESP ability', in J.D. Morris, W.G. Roll, and R.L. Morris (eds.), Research in Parapsychology 1976, Metuchen, N.J., Scarecrow Press, 1977, 122-125.

Dick J. Bierman
RIPP

Alexanderkade 1
1018 CH Amsterdam
The Netherlands

APPENDIX I-A: Raw Data Amsterdam series.

DATE	TIME	SS	Anal.	FB+INST N=1024	NFB+INS N=1024	FB+PRE N=473	NFB+PRE N=472
30/03	11h40	D	DJB	1	3	1	2
30/03	11h45	J	DJB	5	5	1	1
30/03	15h33	J	JMH	0	7	2	1
30/03	15h39	D	JMH	4	4	2	2
06/04	09h55	J	JMH	5	5	5	0
06/04	10h04	D	JMH	2	1	2	4
06/04	11h20	J	DJB	3	4	1	3
06/04	11h25	D	DJB	2	1	4	4
08/04	20h40	D	DJB	3	4	4	1
08/04	21h03	J	DJB	6	3	1	1
08/04	22h23	J	JMH	2	7	0	2
08/04	22h35	D	JMH	5	6	2	1
11/05	09h38	J	JMH	6	4	0	4
11/05	09h42	D	JMH	4	4	1	2
18/05	10h21	D	DJB	6	2	1	2
18/05	10h38	J	DJB	6	0	0	0
18/05	11h58	D	DJB	3	1	1	1
18/05	12h19	J	DJB	5	6	2	3
18/05	13h07	D	JMH	7	2	2	5
10/05	13h25	J	JMH	3	4	0	2
29/06	09h22	D	DJB	5	4	3	2
29/06	09h27	J	DJB	7	6	4	4
29/06	12h12	D	JMH	8	0	1	0
29/06	12h20	J	JMH	3	5	0	0
06/07	09h05	D	DJB	5	5	2	2
06/07	09h14	J	DJB	3	4	1	1
06/07	09h21	D	JMH	4	5	1	2
06/07	12h51	J	JMH	7	6	2	2
06/07	13h17	D	JMH	7	4	0	1
06/07	13h49	D	DJB	3	2	2	0
12/07	09h09	J	JMH	8	5	1	1
13/07	09h45	J	DJB	4	2	0	1
13/07	09h50	D	DJB	6	3	3	2
13/07	09h55	J	DBJ	4	2	0	1
13/07	10h13	D	JMH	4	3	3	2
13/07	11h27	J	JMH	6	2	2	1
13/07	11h32	D	JMH	4	5	1	0
13/07	12h20	J	JMH	4	4	2	1
13/07	12h28	D	DJB	3	2	1	2
20/07	10h13	J	DJB	2	2	1	0
20/07	10h21	D	DJB	2	6	1	0
20/07	13h00	J	DJB	4	0	1	2
20/07	13h44	D	JMH	3	2	1	3
20/07	17h14	J	JMH	4	2	1	3
20/07	17h34	D	JMH	5	5	1	1
20/07	17h53	D	DJB	1	6	1	1
23/07	04h04	J	JMH	1	4	1	2
23/07	04/27	J	DJB	6	5	1	1

APPENDIX I-B: Raw Data Bristol Series.

DATE	TIME	SS	Anal.	FB+INST N=2048	NFB+INS N=2048	FB+PRE N=512	NFB+PRE N=512
08/11	17h15	E	DJB	12	10	0	5
08/11	17h30	E	JMH	10	12	1	0
22/11	14h45	E	JMH	9	6	5	0
22/11	15h45	E	DJB	7	6	4	4
22/11	18h00	E	DJB	12	9	2	5
30/11	12h00	E	DJB	11	12	4	3
30/11	16h23	E	DJB	8	5	1	1
03/12	12h35	E	JMH	6	5	2	0
04/12	11h25	E	JMH	12	5	2	1
05/12	17h13	E	JMH	9	7	3	3
05/12	18h15	E	JMH	6	12	4	3
06/12	13h11	E	DJB	9	10	3	3
06/12	16h50	E	DJB	6	6	2	0
06/12	18h00	E	DJB	14	5	2	1
16/12	16h05	E	DJB	7	12	2	5
16/12	18h05	E	JMH	8	8	2	1
22/12	12h00	E	JMH	11	8	4	2
22/12	14h30	E	JMH	11	9	2	3
04/01	17h20	E	JMH	4	10	5	1
04/01	17h30	E	JMH	12	5	2	2
07/01	17h30	E	DJB	11	6	3	3
11/01	17h10	E	DJB	4	9	1	1
11/01	18h20	E	JMH	5	6	2	4
13/01	14h20	E	JMH	9	9	0	0
13/01	17h00	E	DJB	8	7	2	3
13/01	18h30	E	DJB	8	4	0	3
15/01	14h50	E	JMH	9	4	5	1
15/01	17h35	E	JMH	10	7	4	4
16/01	18h18	E	DJB	10	11	2	0
16/01	18h45	E	DJB	17	6	1	0
17/01	14h00	E	DJB	9	5	2	1
17/01	18h15	E	JMH	8	4	3	0
17/01	18h45	E	DJB	6	9	2	0

APPENDIX II : Formal evaluation of hypothesis using Chi-square.

AMSTERDAM SERIES

Hypothesis 1: Feedback effect.

	Hits	Misses
FEEDBACK	271	71585
NO-FEEDBACK	253	71555

Chi-square=.5422 n.s.

Hypothesis 2: Analyzer effect (pooled FB+NFB).

	Hits	Misses
ANALYZER DJB	247	71585
ANALYZER JMH	277	71555

Chi-square=1.61 n.s.

(only NFB)

	Hits	Misses
ANALYZER DJB	115	35904
ANALYZER JMH	138	35881

Chi-square=1.92 n.s.

Hypothesis 3: Instantaneous vs. Prerecording.

	Hits	Misses
INSTANTANEOUS	357	97929
PRERECORDED	149	45211

Chi-square=2.254 p>.10

BRISTOL SERIES

Hypothesis 1: Feedback effect.

	Hits	Misses
FEEDBACK	377	84103
NO-FEEDBACK	314	84166

Chi-square=5.586 p<.02

Hypothesis 2: Analyzer effect (pooled FB+NFB).

	Hits	Misses
ANALYZER DJB	362	86678
ANALYZER JMH	329	81591

Chi-square=.18 n.s

(only NFB)

	Hits	Misses
ANALYZER DJB	170	43350
ANALYZER JMH	144	40816

Chi-square=.77 n.s.

EUR.J.PARAPSYCHOL., V.5, 1985

DOES IT MAKE SENSE TO BE GENERALLY SKEPTIC TOWARDS SCIENTIFIC DATA ?

Ulrich Timm
Institut für Grenzgebiete der Psychologie und Psychohygiene

I

Piet Hein Hoebens, the late representative of the Dutch section of the CSICOP (Mr. Hoebens passed away on 22/10/1984), has pointed in several articles to inaccuracies in certain publications on the sensitive Gerard Croiset. In so doing he has earned the reputation of a thorough and competent critic of parapsychology. Moreover, unlike other 'psi skeptics', he has proved himself to be relatively cooperative and communicative with the result that the 'Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie' was the first parapsychological magazine to publish contributions from him. In my opinion special attention should be given to his paper 'The Legitimacy of Unbelief' (1982), in which he tries to formulate his standpoint regarding psi phenomena as generally as possible and to substantiate this according to the rules of philosophy of science. He describes this standpoint as being one of 'liberal' or 'soft-line' skepticism. Elsewhere (1980) he also talks of a 'polite disbelief'.

This paper is a slightly abridged and modified version of the German original (Timm, 1982)

Of course opinions may differ as to whether a critic's behaviour in practice can be termed as 'polite' or 'impolite', 'soft' or 'hard'. However, as far as theory is concerned, Hoebens has drawn a clear line between his undogmatic liberal skepticism and dogmatic a priori skepticism: for him there are no a priori judgements on the existence of psi phenomena. As far as he is concerned, the well-known rule that 'something cannot exist which must not exist' is not acceptable. Instead of this he puts foreward the 'fallibilistic' hypothesis (i.e. which can be falsified in principle) that 'Nature does not allow psi'. And he sees it as the liberal skeptic's duty extensively 'to explore the possibilities of the model which is based on the initial assumption that psi is non-existent'. When this assumption is employed it is of course no surprise that so far the liberal skeptic does not see any empirical support for the falsification of this hypothesis and thus his manner of judging concrete empirical material is little different to that of the dogmatic skeptic.

Nonetheless the distinction between liberal and dogmatic skepticism is important: the dogmatic skeptic, who declares empirical data such as those in parapsychology as a priori possible or impossible, true or false, is obviously centuries behind in his understanding of science and can thus be eliminated as partner for scientific discussion. On the other hand, the liberal skeptic does not make his hypotheses absolute - however rigidly he may advocate them. He admits that there is the possibility that they may be false and that counter-hypotheses might be correct. Only this kind of skepticism is acceptable or 'legitimate' in view of the present-day philosophy of science. Thus its 'liberality' does not represent a special qualification, but a minimal demand in empirical scientific standards.

However, in the case of most empirical problems there are a great many possible hypotheses which are all equally admissible or 'legitimate' in a formal sense. Therefore the legitimacy of a hypothesis such as 'Nature does not allow psi' does not state whether this hypothesis is correct or at least superior to other hypotheses. And it is of no use if it has not been falsified up till now, as long as the alternative hypotheses have not been falsified either. Hoebens sees this problem, too, and tries to approach it by having recourse to Lakatos' theory of science. This shall be examined more closely in the following.

II

According to Hoebens, the controversy between parapsychologists and skeptics is in effect a contest between two 'mutually incompatible explanatory models' both of which are produced in order to explain the same data.

Parapsychologists, says Hoebens, explain the empirical data as the effect of an unknown factor 'psi'. The skeptic, according to Hoebens, explains them just as well by reducing them to mistakes, which falsify the data in question. As sources of error he quotes the same possibilities of error, deception, and fraud which are also discussed every now and again by parapsychologists (e.g. by myself, 1981). Among these, fraud by the parapsychological experimenter or reporter plays a great role. This is inevitable since nowadays many experimental settings are secured to such an extent that other sources of error are almost impossible.

The contest between the two models is then described by Hoebens in the terminology of the neo-falsificationist Lakatos (1978), according to which both of the two parties may shield his model against 'premature falsification' by means of any auxiliary hypotheses of his choice as long as these produce satisfactory predictive results concerning future data. Finally Hoebens (without committing himself completely) implies that in this match the skeptic's model apparently stands up to the test better than that of the parapsychologists. For example, it allows the successful prediction of more and more new exposures of errors and fraud, while the psi model may not predict anything at all due to the much discussed instability of psi effects.

Without doubt, there is much to be criticized in that presentation. Among other things one could question Lakatos' theory of 'scientific research programmes' and refer to competing theories of other philosophers of science. However, this would hardly change anything since the result would be that the model contest would simply be described in other terms. In my opinion the really critical point is that those who take that position confuse two different scientific problem levels. On the first level a problem is located which scarcely plays a role in usual epistemological discussions, i.e. the problem of reliability or authenticity and non-falsification of scientific data and protocols (note 1). If one has reason to question the reliability of data one can of course devise hypotheses and models such as Hoebens does. But parapsychologists do this too - and not just by introducing a meaningless 'psi factor' which is completely out of place on this level. Instead of this they discuss exactly as he does the various possibilities of error and fraud and come to a conclusion which is much more complex and predictive than his, i.e. that these effects really explain the data partially but not totally. Under these circumstances, however, an independent second problem poses itself, which is located on a completely different level: namely, the usual problem to describe the structure of those data, which have been ascertained as reliable on the first level, with the help of some kind of scientific theory, the predictive power of which can then be examined on its own. On this level it is true that parapsychologists are not very advanced, but this cannot at all affect their statements about the reliability of data since these simply relate to a completely different level.

After this rejection of the above somewhat too simplistic treatment of the problem, the model contest between skeptics and parapsychologists obviously reduces itself to questions concerning the reliability of data. Questions of parapsychological theory formation would only play a part for those skeptics who - without doubting the facts - discuss 'classic' physical theories for psi phenomena. But most skeptics gave up this attempt a long time ago. The shift in the discussion to the problem of reliability of data represents an unusual but legitimate running fight. It relieves the skeptic from his duty to produce some kind of theoretical hypothesis for parapsychological data, but at the same time robs him of his right to include the theoretical attempts of parapsychologists directly in his discussion. In any case, polemic remarks on the ominous 'psi factor' which explains everything and nothing, are out of place here. Moreover, it should be born in mind that most serious parapsychologists do not use the word 'psi' in the explanatory sense, but only to mark some as yet unexplained empirical phenomena, just the 'psi phenomena'.

However, regarding the question of data reliability - the only legitimate remaining - the argumentation of parapsychologists particularly within the framework of Lakatos' scientific research programme should be clearly superior to that of the skeptics. If it is really a question of shielding a successful model "against premature falsification", the Lakatosian will have difficulties in justifying the replacement of the successful model of a cooperative 'scientific community' - showing an at least passably reliable interpersonal exchange of data - by a completely novel model, according to which every scientist must credit all other scientists with literally any kind of (conscious or unconscious) data manipulation. Rather according to Lakatos' view of protecting a current model, the Lakatosian should produce numerous findings which are consistent with the accepted model but not consistent with the skeptic's new model, and as a result he will reject the latter as falsified. In support of his stance, he could even bring forward the fact that general scientific progress has been made despite numerous scientific errors, cases of deception and fraud, a fact which requires that in science the circulation of correct information markedly exceeds that of misinformation.

Here some skeptics would probably object by saying that they had never maintained that there is a general falsification of data in science, but that this phenomenon is limited to certain 'pseudo sciences' - such as parapsychology. This limitation would of course represent an unequalled example of judging by two different standards unless they could produce convincing empirical evidence for it. Usually they cannot. All they can do is to allude to individual cases in which parapsychologists were found guilty of falsifying data or where this at least seems possible. They could also quote unintentional experimental errors, incorrect statistical evaluation, misinterpretations caused by individual expectations of the experimenter, etc., as well as an undefinite number of cases in which scientific investigators were duped by fraudulent subjects. However, an enumeration of these does not give the slightest evidence that such sources of error are more common in parapsychology than in other sciences. And it is this that must be demonstrated to empirically support the hypothesis that reliability of data is particularly insufficient in parapsychology.

An in-depth examination of the problem of data falsification in various sciences should be based on comprehensive empirical material and should make use of all the resources available from psychology, criminology, psychiatry, etc.. As yet the skeptics have not been able to submit such an investigation (which is, moreover, only significant as long as one assumes that the data for this investigation itself are reliable). True, some skeptics (and Hoebens (1980, 1981) is one of them) deserve recognition for exposing cases of falsification in parapsychology. However, there is as yet no reason to assume that these are not to be found in other sciences if one takes a closer look. In fact, it seems that deception - of oneself or of others - is much more frequent in science in general than its sacrosanct image implies. I myself have (1981) argued the following: "Moreover, scientists are on average, neither in their moral qualities nor in their criticism of self and others, considerably superior to other human beings and they are just as liable to fall victim to certain psychological mechanisms of deception, individual or social. Thus, in science one finds the same kinds of deception which occur in everyday life - that is, everything from deliberate fraud to naive credulity and finally misinterpretation of results caused by individual motivation".

However, the fascinating thing about the development of science in general seems to be that despite this and other obstacles, it has lead to a definite progress in knowledge even if this process cannot be described by a monotonically rising function, but rather according to the scheme of 'two steps forward and one step back'. As yet we have had no reason to postulate an exception to this general rule for parapsychology, at least not in such cases where parapsychology is practised by competent scientists in the form of empirical-experimental investigations.

III

I am completely aware that at least some skeptics will reject the above deliberations as pure conjecture and begin by quoting a class of alleged hard facts with which, as circumstantial evidence, the skeptics apparently hope to convict the parapsychologists, as the latter have neither confessed to general error nor general fraud. It is certain inconsistencies within parapsychological data, which according to the skeptics' interpretation, are said to give almost conclusive evidence of their being falsified in one way or another. At the top of this list is, of course, the 'non-repeatability' of psi results to which Hoebens and others consequently pay special attention. They seem to credit parapsychologists with little interest in this topic, particularly me, since I once declared this problem to be "endgultig gestorben" for me (Timm, 1980). I do in fact take the repeatability problem so seriously that some time ago I subjected it to a reformulation. I had the impression that most authors treat it in a way which is unsatisfactory, both epistemologically and methodologically. There are not even unambiguous definitions for the meaning of 'repeatability of an experiment' or 'replicability of experimental results'. It is thus no wonder that various parapsychologists come to a completely different conclusion on the repeatability of psi results which can easily be learned from a review by Hövelmann (1983). It appears that the idea of repeatability has different meanings for different authors which can lead to considerable misunderstanding, especially in the discussion between parapsychologists and skeptics.

A thoroughly approach must (as I pointed out in 1980) at least take the following points in consideration:
1) There is the logical-ontological axiom of a basic law of causation in nature according to which the same events must always occur given the same initial conditions. It need not be discussed at this point whether this deterministic axiom is absolutely valid (e.g. in quantum physics).
2) From this axiom the general postulate that experimental results must be replicable can only be deduced under a specific presupposition. This presupposition is that in the respective experiments the initial conditions themselves can be replicated exactly at any given time.
3) However, exactly this presupposition is almost impossible in experiments with living organisms, since the latter always vary inter-individually and fluctuate intra-individually in their attributes. Apart from that they react to uncontrollable situational and social conditions of an experiment, which additionally increases the interexperimental variance of the results. Therefore, for biological, psychological, parapsychological and generally (because of the sampling error) for all statistical experiments only a limited ('stochastic') replicability can be expected.
4) The degree of stochastic replicability depends on the extent of the factors mentioned. If psi performances vary considerably because of their natural variability and their great perturbability then one can only expect a small degree of replicability for them.
5) In statistical experiments, to which also belong the usual quantitative psi experiments, the stochastic replicability can be considered as ensured, if many experiments have been conducted, which cover the inter-experimental variability in a representative manner, and if their combined result is statistically significant (note 2). However, the required combination of all psi experiments (or at least of a representative sample of them) already is necessary for pure statistical reasons since separate significance tests of selected experiments (or parts of experiments) may result in serious statistical selection errors (see Timm, 1983).
6) The meaningless demand for strict replicability within a set of statistical psi experiments can thus be replaced by the meaningful

(and anyway statistically required) demand for a significant combined result of these experiments.

The above considerations show that I have taken a serious look at the replicability problem and am even prepared to draw from it extremely grave consequences for parapsychology. It is therefore not at all my intention to evade a further question on this problem, which is discussed by Hoebens as well as by other skeptics: if positive results in psi experiments which have been gained by certain investigators (e.g., parapsychologists) cannot be repeated by other investigators (e.g., skeptics), is then this data inconsistency useful as evidence of any (conscious or unconscious) manipulation of the first investigators? The answer can only be: of course this is a possible implication, but its practical usefulness is so limited that from the 'criminological-standpoint' one can only use it in exceptional circumstances and even then never alone! Skeptics should not make the mistake to believe that non-repeatability could verify their hypotheses by the same degree of certainty, by which repeatability could falsify them. The situation is like the use of an alibi in criminology: an unassailable alibi (repeatability) may be adequate proof for the innocence of the accused (non-falsification of psi data); but the lack of an alibi normally does not prove anything, especially in situations (biological-psychological domain) in which the existence of an alibi cannot be at all expected due to the prevailing circumstances (tendency toward instability of characteristics).

An objective and detailed examination of the problem raised here must submit the whole (or at least a representative cross-section) of the available experimental material to a careful statistical analysis. Only in this way can one decide whether the variability of these data goes back mainly to differences between various experimenters (and thus - if not to an experimenter effect - perhaps to a manipulatory influence), or to fluctuations within the results of one and the same experimenter. The latter finding can stem from differences between the results of various experiments, various subjects or repeated testings of individual subjects. Without wishing to anticipate more exact analyses we can say that only a fraction of the variability of data can be ascribed to the change of experimenters and that the three other factors also play an important role. The most interesting thing seems to be the general fluctuation of individual psi performances over time - not excepting those of high scoring subjects.

Normally one does not expect such differentiated data structures in fraudulent set-ups or in systematic evaluation errors. In particular, one would not expect the imposing palette of 'special effects' which partially counter usual fraudulent intentions (e.g. psi missing as well as unusual patterns of psi performances - as the Rhinean 'quarter distribution') or which only appear in the case of a completely different kind of hit scoring (e.g. consistent missing, displacement effect). In addition there are the correlations between psi performances and other variables, the artificial production of which is not at all so easy. Moreover, the basic assumption that skeptics would never gain significant results nor confirm these by inspection is not correct. It is well-known that in this way some skeptics have been 'converted'. In investigations on the sheep-goat effect - contrary to a widely held opinion - the skeptic subjects often show a significant result, too, though in the form of psi missing.

All these effects are objectively contained in the parapsychological data, whereby the reduced repeatability becomes of much smaller significance than the skeptics seem to believe, as an analysis which is empirically useful must take all available features into account equally and must not take a single one out of its context. However, I admit that there are sometimes data anomalies which so obviously depart from the norm that they could indeed be useful as indicating some kind of falsification. Proceeding from the 'normal' rarity and instability of psi phenomena I earlier (1981) stated the following heuristic rule: "Reports on extremely high or extremely stable psi performances contradict the general experience and, with increased probability, allow the conclusion that it could be - at least partially - a case of error, deception, or fraud". As a consequence, I (Timm, 1983) modified the demand for a combined evaluation of as many psi experiments as possible in such a way that, as a precaution, all results which arouse suspicion (according to this rule) must be eliminated. If the skeptics also confined themselves to such a - risky, but empirically well substantiated - manner of argumentation, the dialogue with them would be much easier than it is to date.

IV

Finally, I would like to summarize my views as follows: The hypotheses expressed by the skeptics are indeed just as 'legitimate' as numerous other hypotheses which can be advanced to explain psi phenomena. However, science only results when hypotheses are tested empirically and evaluated in an efficient manner. Consequently, if one examines the hypothesis that all parapsychological findings are falsifications (in the widest sense of the word), then one can only adjudge to that hypothesis a minimal inductive probability, even when applying critical standards. This probability is so minimal that from a strategical point of view it is useless to follow up this hypothesis. It is in any case more advantageous for scientific progress to subject the more probable hypotheses to a thorough empirical test, as parapsychologists do. Should it yet turn out that the most improbable of all hypotheses is right, then this research will exhaust itself - for one cannot in the long run build up a scientific theory on falsifications!

NOTES

1) Here in general all descriptive statements concerning empirical findings of a science are termed as 'scientific data and protocols'. In the concrete case one assumes that these are formulated in an intersubjective observation language which is understood by the skeptics in the same sense as it is by the parapsychologists (which is fully the case). In such a language one can describe apart from quantitative data (e.g. number of hits in psi experiments) also qualitative facts, for example that in a certain experiment a subject attained knowledge of events which a) where so far away from him or her spatially, temporally, or physically (e.g. due to screening) that one can exclude the transference of information by known physical processes; b) which could not have been known to the subject beforehand since they were only produced during the experiment by a random event generator; c) which, according to a generally accepted method of statistical analysis, cannot be traced back to chance guessing etc., because of its high significance. Apart from observation statements the protocol may also contain simple and unambiguous conclusions in as far as they do not go beyond the framework of the scientific world-picture accepted until then (e.g. "sensory contact by means of known sense organs was excluded"; on the other hand not: "there was a telepathic contact".).
When discussing the reliability of scientific data here it is exclusively a question of whether empirical facts have been correctly or incorrectly recorded and documented. Causes of possible inaccuracies could be both lacking observation and control as well as

subsequent errors and falsifications while processing and documenting. On the other hand basic epistemological arguments against the objectivity, independence of theory, and accuracy of measurement of scientific data do not play a part in the discussion given in this paper.
2) To avoid that the combined result is unreasonably reduced by the great variability of single results, I suggest special combination methods weighting the single results according to their size. I also show that the usual addition of hits in psi experiments is one of the most unfavourable methods of combining results (Timm, 1983).

ABSTRACT

Recently P.H. Hoebens (1982) discussed two 'incompatible' explanatory models of psi phenomena, one adopted by parapsychologists and the other adopted by the skeptics. In his view both have the characteristics of a research programme as proposed by Lakatos, and since neither of the two is clearly superior, they can both be regarded as equally justified. Nevertheless, a careful analysis reveals that the model proposed by the skeptics expresses general doubts about the reliability of data in science or at least in parapsychology. Yet this model is a decidedly revolutionary one which need to demonstrate its advantages compared to established and relatively successful models about the reliability of scientific data and not vice versa. To date the 'hard core' of the established theories do not include the assumption that faulty data do not exist at all but implies that faulty data do not occur sufficiently frequently and systematically to substantially distort the total picture of all the data of an empirical science. As a consequence the assertion that data are at fault has to be empirically justified for every single case and should not be a priori based on an as yet insufficient theoretical explanation of the data, as is usually done by skeptics with regard to parapsychology. The empirical evidence presented by the skeptics to date is in no way sufficient to doubt the reliability of the entire data in empirical-experimental parapsychology.

REFERENCES

Hoebens, P.H. 'Vom Lob der Genauigkeit in der Parapsychologie', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1980, 22, 225-234.

Hoebens, P.H. 'Gerard Croiset: Investigation of the Mozart of 'Psychic Sleuths'', Skeptical Inquirer, 1981, 6, 17-28.

Hoebens, P.H. 'Die Legitimitat des Unglaubens', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1982, 24, 61-73.

Hövelmann, G.H. 'Zum Problem der Wiederholbarkeit parapsychologischer Experimente', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1983, 25, 29-54.

Lakatos, I. 'The methodology of scientific research programmes'. In 'Philosophical Papers', Cambridge, University Press, 1978, 8-101.

Timm, U. 'Wie unwissenschaftlich ist eigentlich die Parapsychologie?', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1980, 22, 103-116.

Timm, U. 'Die Unwahrscheinlichkeitsregel - oder: Konnen Psi-resulate 'zu gut' sein?', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1981, 23, 207-217.

Timm, U. 'Ist ein genereller Skeptizismus gegenuber wissenschaftlichen Daten sinnvoll?', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1982, 24, 167-182.

Timm, U. 'Statistische Selektionsfehler in der Parapsychologie und in anderen empirischen Wissenschaften', Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1983, 25, 195-229.

Dr. U. Timm,
Institüt fur Grenzgebiete der Psychologie und Psychohygiene,
D-7800 Freiburg i. Br.,
Eichhalde 12,
B.R.D.

LETTERS TO THE EDITOR

From Dick J. Bierman: Towards a better methodology in research with psychics

The article 'The potential paranormal value of statements of psychics acquired under feedback conditions' by Boerenkamp that appeared in the E.J.P., 5, 2, 1984 seems to describe a promising tool that might be used to study paranormal phenomena in the field and under natural conditions.

However the author fails to indicate a major potential criticism that might arise if the analysis based upon this method yields significant results. The protocols that are studied were acquired in a face to face situation. Therefore the written protocols do not reflect the total information flow from and to the psychic. It is well known that non-verbal communication plays an important role and in fact many stage performers use this type of information in their 'psychic' acts. Since, as far as I can see, there is nothing in the method described by Boerenkamp that would invalidate this type of criticism, it must be concluded that his work could never have yielded data from which conclusions could be drawn with regard to the psi hypothesis.

There are two possible extensions to the methodology that I should like to propose:

1. The first one is to videotape the session, especially the sitter, and use the video protocol as a basis for the judging procedure. This would enable the judges to take into account the aspects of non-verbal

communication that are clearly visible like nodding, rubbing hands etc.. However one should realize that minor phenomena like the very rapid changes in facial expressions have been shown to form a (subliminal) aspect of non-verbal communication. These fast phenomena can not be recorded on video (high speed film is better but much more expensive).

2. A better approach therefore would be to separate the sitter and the psychic in order to avoid any visual contact while maintaining other more controllable ways of communication. Audiotapes from such sessions could then be used for further analyses along the lines sketched by Boerenkamp. Judges should try to incorporate information hidden in intonation, hesitations and the like. A further step in this direction is to have the sitter and psychic communicate by electronic mail. Such a situation has the drawback of being quite different from the natural setting but retains the aspect of (slightly delayed) feedback. The advantage, of course, is that the room for non-psi explanations, if significant results are obtained, is drastically reduced. Also the on-line storage of the protocol is an advantage.

The biggest advantage however is that this set-up can be used in conjunction with taking a 'thinking aloud' protocol of the 'psychic experiences' and their statements. Although the request to think aloud, while communicating with the sitter through a teletype-link, might interfere with the normal way of experiencing the impressions by the psychic, it nevertheless might shed light on (for instance) the associational networks that the psychic 'uses' between impressions and statements.

Thus our efforts to improve our methodology would be geared towards a better understanding of the phenomena instead of towards another 'proof' that the phenomena are real.

Dick J. Bierman
RIPP
Alexanderkade 1
1018 CH Amsterdam
The Netherlands

(received December 1984)

From J. de Boer: Randomness and Probabilistic Predictor Programs

In the last years two papers on the concept of Probabilistic Predictor Programs appeared in this journal. Tart and Dronek (1982) introduced the concept in connection with the randomness or non-randomness of a sequence of numbers or items (the target sequence) and the responses (calls) of a percipient as to what he believes the order of the target sequence is when he gets feedback of target identity after each call. The authors developed an algorithm, called Probabilistic Predictor Program (PPP), as a test of the degree to which non-randomness of the target sequence might allow a mathematical inference strategy by the percipient to account for results. If the results (number of 'hits') attainable by this test are considerably lower than those actually obtained in the experiment, it is considered legitimate to argue that ESP was operating. According to the authors, in this way a biased target sequence need not be totally discarded and their PPP or some superior version of it is proposed as the standard measure of predictability. Vassy (1984) improved the test to PPP-B (what does B mean here?) and endorsed the proposal for standardization.

Some objections may be raised here. First it is possible that a percipient scores considerably higher than according to the PPP of Tart and Dronek, but not higher than according to the PPP-B of Vassy. Then the conclusion would be that ESP is operating or not operating contingent upon whether the PPP of Tart and Dronek or the PPP-B of Vassy is taken as the standard. So the conclusion depends on whatever Probabilistic Predictor Program happens to be used as standard and hence the conclusion may change when an improved version of the standard is introduced.

A second objection could be lodged against the use of biased target sequences. Tart and Dronek state correctly that randomness of the target sequence is necessary for the standard statistical tests to be valid and this is enough to reject a biased sequence.

Finally both objections are met if one knows beforehand that the target sequence may be considered random, thus eliminating the need for PPPs. But the only way to know beforehand that a sequence is random, is to produce it according to a certain well-defined algorithm, a so-called pseudo-random number generator, an RNG. This sounds paradoxical but such a sequence can - when not too short and under certain restrictions on the algorithm - be proved to be practically random. It is not the place here to go further into this highly sophisticated mathematical subject, which moreover is still in development, but the interested reader can find an excellent overview of this topic in Chapter 3 of Knuth (1981). One might perhaps think that a percipient can use the same algorithm, thereby scoring 100%, but such an algorithm is far too complex for the human mind. On the other hand it is perfectly fitted to production by a (micro) computer.

It is clear that such a method of random number generation is far superior to other methods, as e.g. shuffling of cards, because of its guaranteed randomness, reproducibility and ease of generation. Tart and Dronek state that such an RNG is rarely used in parapsychological research but the arguments given above suffice in my opinion to propose its general use, and possible standardization of one specific RNG. By choosing different starting numbers different sequences can be obtained. There is no danger, in case of the possible replacement of the standard after some time, that conclusions of former experiments are invalidated because all standardized sequences are random and replacement will only be proposed on practical grounds, e.g. word size of the (micro)computer or computation speed. Further details could be discussed when the proposal meets with sufficient approval.

Knuth, D.E. 'The art of computer programming', Vol.2: Seminumerical algorithms, Second edition, Addison-Wesley Publishing Company, 1981.

Tart, C.T. and Dronek, E. 'Mathematical inference strategies versus psi: Initial explorations with the Probabilistic Predictor Program', E.J.P., 4, 3, 1982, 325-355.

Vassy, Z. 'Improvement of the 'Probabilistic Predictor Program' of Tart and Dronek for testing random target generators', E.J.P., 5, 3, 1984, 203-219.

J. de Boer,
Zevenenderdrift 40,
1251 RC Laren,
The Netherlands.

(received June 1985)

ERRATUM

We regret a number of serious printing errors in the November 1984 issue of this journal.

p.194,line 5: the words 'debates around' are missing between 'continuing' and 'parapsychology'.

p.195,line 7 from bottom must read: 'Others of his activities in support of reasonable and responsible parapsychology will be made known in due course, and their importance to parapsychology will become more evident as time goes by'.

p.196,line 7: it must read 'for the record'.

p.196,line 13: 'misinterpretation' should read 'misrepresentation'.

p.196,2nd paragraph,line 1 should read: 'attach much importance'.

p.197,3rd paragraph,line 2: must read 'should have a very careful look at'.

p.198,last paragraph,line 2: must read 'and responsible critics alike is that only rational and fair-minded dialogue and cooperation'.

p.198,last paragraph,line 5 from bottom should read: 'superiority of one's respective position, which we frequently encounter on'.

p.200, 2nd reference should read:

Clark, J. 'The rational inquirer', Fate, 1983, 36, 93-100.

and:

Frazier, K. 'Parapsychologists, critics agree to consensus statement', Skeptical Inquirer, 1984, 7, 4-6.

p.285,line 6 from bottom should read: 'robust and repeatable' instead of 'robust and hidden'.

p.289,line 2 must be: 'and interpretable relative to the purposes'.

p.289,2nd paragraph,line 2 should read: 'in pursuing one's

p.296,line 12: 'methodical' must be 'methodological'.

p.297,line 3: 'immediate' must be 'intermediate'.

p.299,in the right column of the schema replace S_1 by S_2.

p.303,line 1: the name must be Capra.

Editorial consultants: Volume 5, 1983-1985.

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.

Dr. Eberhard Bauer (University of Freiburg, Germany); Dr. J. Beloff (University of Edinburgh, U.K.); Dr. R.S. Broughton (F.R.N.M., U.S.A.); Drs. G. Camfferman (University of Utrecht, The Netherlands); Dr. Hoyt Edge (Rollins College, U.S.A.); Dr. E. Haraldsson (University of Iceland); Dr. J. Houtkooper (University of Amsterdam, The Netherlands); Dr. E.F. Kelly (Spring Creek Institute, U.S.A.); Dr. B. Loefqvist (University of Lund, Sweden); Dr. W. von Lucadou (University of Freiburg, Germany); Dr. Op 't Hof (University of Amsterdam, The Netherlands); Dr. J. Palmer (F.R.N.M., U.S.A.); Dr. T. Pinch (University of Bath, U.K.); Prof. G.W. Smith (University of Lund, Sweden); D.H. Weiner (F.R.N.M., U.S.A.); R.A. White.